DOCUMENT RESUME

ED 032 584 CG 004 485

By-Lieberman, J. Nina

Personality Traits in Adolescents: An Investigation of Playfulness-Nonplayfulness in the High-School Setting.

City Univ. of New York, N.Y. Brooklyn Coll.

Spons Agency-National Inst. of Mental Health (DHEW), Bethesda, Md.

Pub Date Dec 67

Note-97p.

EDRS Price MF-$0.50 HC-$4.95

Descriptors-*Adolescents, *Childhood Needs, High School Students, Personality Assessment, *Personality Studies, *Student Needs, Student Teacher Relationship, Teacher Response

The Specific aim of the present project was to establish criteria for playfulness and nonplayfulness in adolescents and to develop measuring instruments for these behavior dimensions. It was hypothesized that teachers were able to rate adolescents on these traits as manifested in the high school classroom. Questionnaires were completed by 115 junior high and high school teachers. On this instrument, ten subscales showed satisfactory reliability and validity coefficients. Two distinct factors emerged, and on the basis of their loadings on component traits of playfulness and nonplayfulness, were labeled "social-emotional" and "academic" playfulness and nonplayfulness. Conclusions drawn from the data include: (1) playfulness and nonplayfulness are identifiable in the high school classroom and are complementary ends on a rating scale continuum, and (2) teachers see their students as relatively stable in playfulness-nonplayfulness characteristics over a period of one to two months. (Author/KJ)

ED0 32584

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

N-X

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

PERSONALITY TRAITS IN ADOLESCENTS:
An Investigation of Playfulness-Nonplayfulness in the High-School Setting

J. Nina Lieberman
Brooklyn College of the
City University of New York

CG004485

ACKNOWLEDGEMENTS

I wish to thank the cooperating schools, the teachers who acted as raters, and my colleagues, Dr. Abraham Bernstein and Dr. Jack Campbell, who involved their students as raters in the project, for their wholehearted support and help in this investigation.

The suburban sample was drawn mainly from the Mount Vernon High School and the investigator is indebted to Dr. Norman Eagle, Director of Research, and Mr. E. J. Cipolaro, Principal of the High School, for their understanding and interest. My thanks also go to Mrs. Ethel Wengraf, the Coordinator of Research for Brooklyn College at Ditmas Junior High School in Brooklyn as well as to Rabbi David Eliach, Acting Principal of the Flatbush Yeshivah High School in Brooklyn, who were instrumental in involving teachers at their schools in the project.

This study would not have been possible, of course, without the enthusiasm and cooperation of the teachers and student teachers. I, therefore, want to express my gratitude to Mr. Abraham Carmel, Mr. Martin J. Cohen, Mr. Anthony De Rosa, Mrs. Luba Ellis, Mrs. C. Irene Green, Mr. Amnon Haramati, Mrs. Marian Hurst, Mr. Ronald C. Kawesch, Mrs. Janice D. Morganstern, Mrs. Rose L. Pucci, Miss Maryann Rosa, Mr. Ashley Silverman, Mrs. Harriet Spitzer, Mr. James Yett, and Mr. Morton J. Zerder, teachers at the respective schools, and Mr. Joseph Barone, Miss Myrna Davis, Mr. Leonard Mazin, Miss Marilyn Murray, Miss Linda Podonsky, Mr. Sydney Rosen, and Miss Janet Robbins, student teachers in the New York City system.

I should also like to express my appreciation to Professor Hyman Sardy, who acted as statistical consultant, and whose encouragement and genuine interest were a great help. My special thanks also go to Professor Robert L. Thorndike whose comments on the data analysis and the terminal report were invaluable.

Miss Felice Bernstein assisted in the early part of the data analysis and Miss Judy Garretson was associated with the project during the latter part. They both gave unstintingly and enthusiastically of their time and energy and made many helpful suggestions.

Last but not least, I am indebted to Mrs. Gilda Koeppel whose patience in deciphering the manuscript equalled her competence in typing it.

This study was supported by a Public Health Small Grant MH 13152-01 from the National Institute of Mental Health.

December, 1967
Brooklyn, New York

I - SOME ASPECTS OF THE RELATIONSHIP BETWEEN PLAY AND THOUGHT

Though a rich source of data, play has been neglected as an area in which its specific characteristics could be precisely identified and measured. Some seven years ago, when this investigator first became involved in the analysis of play, the studies in the literature showed two major approaches to interpreting play, namely, educational and clinical, and the emphasis seemed to be on either the global learning aspect or on some specific actions indicative of disturbance.

It does seem equally important, from a developmental viewpoint, to inquire more closely into individual style of play and the carry-over of some of the characteristics of play into later life.

Every so often in investigations of divergent thinking the term "playful" is mentioned as a personality correlate. Getzels and Jackson (1962) found it to be a differentiating trait between creative and intelligent adolescents in their fantasy productions on the TAT. The individual profiles presented by Wallach and Kogan (1965) underline the element of playfulness in their high-creative:high-intelligent youngsters. In our investigations (Lieberman, 1964; 1965) a relationship between playfulness as a quality of play in kindergartners and the divergent thinking factors of ideational fluency, spontaneous flexibility, and originality was found.

Empirical observations, experimental data and theoretical formulations as shown by Lieberman (1966), Piaget (1945), Rogers (1959) and Torrance (1962) lead to the assumption that playfulness as a quality of play survives the age of play and may become a personality trait of the player, and possibly a clue to cognitive style.

While playfulness is commonly ascribed to the creative individual, the behavior need not necessarily be confined to the realm of abstract thinking. Using Guilford's (1956) content areas of intellect--figural, semantic, symbolic and behavioral--we might conjecture that playfulness applies across the board and that if and when encouraged in the individual, it can allow the grown person added scope in his preferred area of functioning or his developed field of competence. It might be worthwhile to cite here Sutton-Smith's (1966) criticism of Piaget's approach to play, in which he maintains that Piaget is unconcerned about the carry-over of some elements of play into cognitive development. Our position here is that while our primary interest is in playfulness as a clue to cognitive style, playfulness in other areas of functioning must also be acknowledged.

In our present investigation we chose to focus on the adolescent because it is the time for consolidation of crucial personality dynamics. Theoretically and empirically, playfulness is part of the adolescent's behavior pattern. Cognitively, the adolescent is capable of "as if" thinking which seems a prerequisite to toying with ideas and concepts as well as seeing remote connections. Emotionally and socially, there is an expanding of the horizon and a loosening of previous bonds. Physically, a new energy reservoir is opened up.

We narrowed down the area of observation to the high-school classroom, because our first concern was to establish whether playfulness is part of classroom behavior and if so, how it can be identified and whether it can be measured. To recognize these behavior tendencies in the classroom and utilize this energy constructively for the learning process was the major rationale for this study. Keeping spontaneity and fun as part of the learning process may be regarded as an important contribution to personality dynamics and educational practice.

The first question that had to be tackled was whether the behavioral correlates of the earlier study at the kindergarten-level could be used in the high-school classroom. The second question was whether, in conceptualizing playfulness, the dimension of nonplayfulness also needed to be spelled out.

The operational definition of playfulness at the kindergarten-level was physical, social and cognitive spontaneity, manifest joy and sense of humor, and a rating instrument constructed along these behavioral indices was used in that study. Although indices for nonplayfulness were also established such as distressed, rigid, perseverative behavior, no attempt was made at the kindergarten-level to incorporate these traits into the instrument. The measurement at that level focused only on the quantity and quality of playfulness itself.

At the adolescent level, however, it was felt that in order to round out the concept and the behavior subsumed under playfulness, a formulation of nonplayfulness was also necessary. It was hoped in this way to remedy what might have been a shortcoming at the kindergarten-level.

The specific aims of the present project were to establish criteria for playfulness (PF) and nonplayfulness (nonPF) in adolescents and to develop measuring instruments for these behavior dimensions. It was hypothesized that behavioral indices for physical, social and cognitive spontaneity, and manifest joy can be identified in adolescents and that teachers are able to rate adolescents along these traits. As a corollary hypothesis, the dimensionality of playfulness was tested. While PF had been found to be unitary in kindergartners, references to the trait at the adolescent level and data from a pilot study suggested either its fragmentation into playfulness (spontaneity), sense of humor and manifest joy or a manifestation in one or the other area of functioning, i.e., either physical, or social, or cognitive.

It had originally been hoped to work with a reformulated PF-Scale based on the Kindergarten-model and the chief investigator's informal observations and interviews with teachers. However, a trial run of that scale with older adolescents indicated the desirability of further empirical work. It was therefore planned to develop and construct a measuring instrument for both playfulness and nonplayfulness in the same manner as had been done for the PF-Scale Form K, i.e., to collect behavioral criteria from teachers and other psychologists and to reformulate the PF-Scale for adolescents on the basis of these data. Furthermore, the reliability of the ratings was to be established by the test-retest method, based on two separate rating sessions, in addition to having inter-rater reliability for each rating session.

Tightening the conceptual framework and the operational definition of PF as well as establishing the validity and reliability of the PF-Scale would then provide a measuring tool for the follow-up study which has as its objective the testing of the relationship of PF-nonPF to divergent thinking as well as assessing the influence of age, sex, teacher personality, and classroom structure on the manifestation of PF-nonPF.

II - DESIGN AND PROCEDURE

In accordance with one of the stated broad objectives of the long-term investigation, namely, to determine whether behavioral indices of playfulness and nonplayfulness are comparable at the various age levels, the design of this investigation first had to establish criteria for playfulness-nonplayfulness at the adolescent level, and, even more specifically, behavioral criteria for high-school students in a regular classroom. The second phase of the investigation was to measure the behavior conceptualized as PF-nonPF.

The procedural steps for the development of the PF-nonPF measure in adolescence were modeled after the investigation at the kindergarten level, with certain modifications and additions.

The sequence of data collection was carried out in five phases as follows:

1. Checking of Criteria for Playfulness (PF) and Nonplayfulness (nonPF) by Investigator.

The chief investigator visited seventeen classes in two private schools in New York City. The subject areas covered were English, Social Studies, History, Mathematics, Chemistry and Physics in grades 9, 10, 11, and 12. The total number of students observed was approximately 300. The schools were chosen because of their predominantly middle-class population which, according to the investigator's conceptualization, would facilitate the observation of PF-nonPF. The schools were atypical in two respects: one, Packer Collegiate Institute, is an all-girls' school, and the other, Flatbush Yeshiva High School, carries a double curriculum, namely, English and Hebrew. However, for the specific behaviors sought, these were not considered factors that would unduly influence the sampling of behavior.

Arrangements for classes to be visited were made through the respective principals of the two schools. The investigator introduced herself to each individual teacher as she was about to visit the class. A brief statement was made to the teacher about the focus of the observations which made it necessary that the investigator sit facing the class. The populations in both schools were well accustomed to "visitations" and there seemed to be no noticeable influence of the investigator's presence on their behavior. When the class session ended the investigator usually had another chat with the teacher which sometimes was more exhaustive than the introductory talk. During this chat, she asked for the teacher's reaction to the proposed investigation and the behavior that it sought to determine. While there was general friendliness about allowing the investigator to visit, two schools of thought emerged even among the small sampling of teachers. Roughly, it could be described as a positive or negative approach to PF-nonPF itself. Translated into the classroom setting, it was seen as either a constructive or a destructive influence on the learning climate. This, in a way, foreshadowed the two clusters of PF-nonPF in teachers' observations of high school students in the standardization sample.

In order to compare individuals' behaviors in different situations, the investigator also followed both a teacher and a group of students into three different class meetings. The investigator used the critical incident technique, i.e., the most obvious and, from her conceptualization of

PF and nonPF, the most pertinent behavior was recorded and later coded for physical, social, and cognitive spontaneity, manifest joy and sense of humor, and their counterparts on the nonPF dimensions, namely, physical, social and cognitive rigidity, manifest distress and tenseness, and matter-of-factness.

2. Validation of Investigator's Criteria against Criteria of Fellow Psychologists.

Four psychologists, three males and one female, whose work as teachers and as clinicians involved them with adolescents, were the sample from which criteria of PF-nonPF in adolescence were collected. Their orientation was chiefly psychoanalytic, but varied in degrees of orthodoxy.

Interviews with the psychologists were conducted at Brooklyn College, either in the psychologists' offices or at the investigator's office. Each interview lasted from about thirty to forty-five minutes. The interviews with fellow psychologists were conducted with a minimum of structure. Initial questions about their conceptualization of PF-nonPF were asked with the request to supplement this with a concrete behavioral example. Where necessary, follow-up questions were added for clarification.

3. Validation of Investigator's Criteria against Teachers' Criteria of PF and nonPF.

A cross-section of high-school teachers were asked to state their concepts of playfulness and nonplayfulness in adolescents in a questionnaire. The Teacher Questionnaire was developed as an open-ended instrument to tap the conceptualization of playfulness and nonplayfulness as seen by JHS and HS teachers.

Six questions pertaining to PF and nonPF formed the body of the questionnaire. The questions went from the more general to the more specific, i.e., concept of PF and nonPF in general, PF and nonPF in adolescents, and then traits and incidents describing playful and nonplayful behavior in a student in the classroom. While the format of the scale tried to give the respondent as much freedom as possible, the use of the term "playful student' and nonplayful student" might have produced a bias toward dichotomizing the behavior.

Two questions aimed at assessing the teacher's own feelings about PF and nonPF in terms of whether it influences classroom learning and if so, how.

Other information provided by the questionnaire was the name, school, and subject matter area of the respondent.

In addition to the questionnaire itself, the Provisional Form of the PF scale-A (without the nonplayful dimension) was distributed. This form had originally been adapted by the investigator from the PF-Scale Form K. After completing the questionnaire, the respondents were asked to comment on the Provisional Scale, and in particular, whether these traits as stated were observable in the classroom. As a concrete example they could use the scale for rating one of their students.

Appendix A gives the Teacher Questionnaire together with the instructions and the Provisional PF-Scale. One hundred and fifteen junior-high-school and high-school teachers completed the open-ended questionnaire. They consisted of 102 regular teachers and 13 student teachers. Of the 102 regular teachers, 7 taught in private schools and were approached through the principals; the remaining 95 were teachers who attended graduate courses at Brooklyn College during the summer of 1966. They, as well as the undergraduate student teachers, were approached through their instructors. The distribution by sex was 46 male and 69 female teachers; by level of teaching, 61 junior high school and 50 high school teachers. Four could not be determined from the questionnaire. The breakdown of teachers by subject matter area was 21 in Social Studies, 20 in Mathematics, 16 in English, 12 in Modern Languages, 11 in Science, 10 in Physical Education, 8 in Home Economics, 5 in Art and one each in the following subjects: Remedial Reading, Speech Improvement, Guidance, Industrial Arts, Distributive Education, Reading and Spanish-Mathematics.

Briefing sessions were held with all of the 115 teachers except for three in one private school at which the information was given out by the principal. Class time lasting from 30 to 40 minutes was allocated by the respective instructors in which the investigator gave the background of her previous work with kindergarten children in that area and explained the reasons for approaching high-school teachers at this phase of the investigation. Some time was also left for the teachers' queries. The questionnaire was then completed in the remaining time. In four classes all of the graduate students present completed the questionnaire, including those that were not teaching at the high school level. (The material from the latter group is being held for analysis at a future time.) In one class, the instructor asked for volunteers, and seven students then went into a separate room with the investigator. Initially there was a feeling of "what's all this about," but once the teachers became involved in the actual completion of the form there was full cooperation, and, on the whole, a positive attitude.

4. Reformulation of PF-Scale and Construction of NonPF-Scale Measure.

While the investigator herself had accumulated extensive records on approximately 300 HS students and had added to these empirical data the conceptualizations of other psychologists, the basis for the final formulation of the measure was the answers obtained on the open-ended questionnaire from the 115 teachers.

This decision was taken after trials run with a Provisional Scale of the PF dimensions adapted for the adolescent from the K-format. The first probe in the Summer of 1965, with graduate education students who were HS teachers, revealed that while the behavior outlined in the Scale gave some information on what the teachers considered a "playful student," other subjects "slipped through." The investigator herself rated one of her own undergraduate classes of late adolescents on the Provisional Scale and gave the students a Self-Scale, tapping the same behavior, to complete. By inspection, correlations between the teacher-observed and the self-rated characteristics were running from the .30's to the .50's. What seemed however even more important than the size of the correlation coefficients, were the comments of the students who maintained that their "playfulness" or manifestation of it was to some extent influenced by the subject matter area and by the teacher.

Inspection of the data collected on 300 students by the investigator during the spring of 1966, which extended the PF-dimension to include also nonPF traits, showed the largest number of recorded behaviors in the area of physical spontaneity, followed by a tie among traits describing sense of humor and cognitive spontaneity and manifest joy, and considerably fewer in social spontaneity. It is interesting to note here also that the number of PF traits mentioned equal those of nonPF in physical and cognitive spontaneity, and are roughly two to one in manifest joy and in a three to one proportion on the PF end in sense of humor and social spontaneity.

The visits in which the same teacher was observed in different classes and the same class with different teachers did not show clear-cut trends, partly because of the variables involved and partly because of the way of assessing them, namely by critical incident, which proved hard to quantify. What could be seen and can also be stated here as the most definitive element is that teacher-personality (or classroom climate), student personality and subject matter area interacted in the overt manifestation of PF-nonPF traits. At the same time, there seemed to be a qualitative constant in the student and it appeared that classroom climate and subject matter area influenced the quantitative element of PF-nonPF. The latter observation was an important consideration in retaining, in the final format of the PF-nonPF measure, a quantitative and qualitative assessment of the traits in the physical, emotional, social and cognitive domains.

Examination of the criteria obtained from the four psychologists suggested the elements of ease and freedom of bodily movement and, its counterpart, inhibition to do so as ingredients of PF-nonPF. In the emotional area, there was a departure from the teacher-culled conceptualization inasmuch as enthusiasm combined with achievement orientation was seen as nonPF and the teaching or classroom situation itself was construed as bringing about nonPF, especially through its emphasis on goal-direction and planning, which would tend to rule out free expressions of joy and other forms of emotional spontaneity. There was agreement, however, on the behavioral critera of sense of humor, and the ingredient of hostility in wit. Psychologists seemed to place little if any emphasis on group-oriêntation as an important part of PF. One mentioned a kind of group-orientation as nonPF, and that some behaviors mentioned as PF especially in the intellectual domain, could only be pursued alone. Two implied that intellectual liveliness and spontaneity were in some way positively associated with PF and one considered day-dreams as a source of PF. As to the kind of work-involvement stipulated as PF-nonPF by teachers, half of the psychologists agreed with the teachers on the erratic vs conscientious dimension. In summary, it might be said that while the psychologists' criteria provided a useful validity check on the situation-spanning traits, the situation-specific traits were considered to give the most accurate picture of the PF-nonPF student. This strengthened the decision to rely almost exclusively on the teacher-stipulated traits supported by the observed congruence between those and the ones found by the investigator and the psychologists.

Support to restrict the formulation of the new PF-nonPF measure to the teacher-criteria came also from the comments made by the 115 JHS and HS teachers after they used the provisional scale to rate one of their students. Although only about half of them made notations on the Provisional-Scale some of the more pertinent critiques were against one-word descriptions, lack of choices being offered, a bias toward PF over nonPF (though the latter was not mentioned in the Provisional-Scale), and the influence on ratings of age, subject matter area, and individual student personality. Bodily expressions of PF were considered significant indices as was sense of humor. Some misgivings were voiced about teacher-bias.

Working with the answers to the teacher-questionnaire, the investigator herself sampled some 25 and arrived at provisional categories that suggested for general concept of PF-nonPF, motor, social, intellectual, emotional manifestations, a combination of both, and a category for "other." For general behavior in a teenager, physical activity, social interaction, imagination, sense of humor and joy were found to be PF dimensions. On the nonPF end were physical rigidity, social aloofness and rigidity, anxiety, intellectual rigor, matter-of-factness (seriousness), and boredom. Categories for the actual PF-nonPF student in the classroom were first seen in terms of categorizing them as "related or unrelated to ongoing work." There was also evidence of positive or negative connotation in the statements relating to PF-nonPF.

Without any prior briefing, a graduate assistant was asked to categorize the answers of a randomly selected sample of 36 questionnaires. Inspection by the chief investigator showed the same patterning as the categories set up by her but while there was general agreement on these categories, the material did not lend itself to the type of coding that would allow for percentage computation of coder agreement. Two sample runs reached from between .40 to .60 agreement. Another way had, therefore, to be found to translate what the teachers had said into an instrument to assess PF-nonPF. Since the main emphasis in the original design was on teacher-agreement and not coder agreement of teachers' responses, it was considered justified to use this criterion as the basis of formulating PF-nonPF dimensions. Moreover, a consistent trend by teachers to start describing an individual student already in answer to the first question of conceptualization and increasingly so when a teenager's PF-nonPF was asked for led to the decision to use, as the basis for the final formulation of the PF-nonPF instrument, answers to questions (3) and (6) on the Teacher-Questionnaire, namely, those describing specifically the playful and nonplayful student in the classroom. Our major interest was, after all, in the behavioral criteria that could be observed in the classroom.

This approach was found to be workable, both from the conceptual and statistical point of view. Quantitative and qualitative dimensions emerged again. For example, actual physical mobility was labelled the quantitative aspect. The energy investment of physical functioning, i.e. alertness versus apathy, was tapped as a qualitative ingredient. This approach was reflected in the phrasing of the questions, namely, "How consistently does the student show spontaneous physical movement and activity in class," and "What degree of energy does the student show in physical activity?". The same approach was used in the phrasing of the four other behavioral dimensions. A special point was made to retain as much as possible of the teachers' phrasing of behavior in establishing the profiles. The questions introducing the scale and the terms at the extreme end were arrived at by agreement among the investigator and the graduate assistant. A final check was carried out by the statistical consultant for logical internal consistencies of the traits, and a trial run with a group of graduate students who were JHS & HS teachers further tightened the profiles. On the whole, the teachers in the pilot study were able to rate their students along the behavior dimensions suggested, and they commented that they found the scales easy to use and appropriate for their groups.

A random sample of 16 questionnaires was also content-analyzed for possible difference in conceptualization between JHS and HS teachers and it was evident that PF-nonPF behavior was seen along the same traits on both levels.

The final form of the PF-nonPF measure consisted of ten subscales. They tapped the following dimensions of PF-nonPF: physical, social, cognitive, emotional, and sense of humor. Each scale was introduced by a question about the student, which was to be answered by rating him on a five-point continuum. The extremes of the continuum were labelled, and a profile of the student at both ends of the scale was also given. The questions alternated in asking either for a quantity or quality assessment of the five areas covered.

Two "ringer" questions, considered not related to PF-nonPF behavior, were added. These asked about the student's achievement-orientation and his physical attractiveness.

A copy of the final form of the scale, the rating instructions as well as a sample trait rating sheet, are included in Appendix B.

5. Standardization of PF-nonPF Measure.

A cross-sectional representative sample of students in grades nine through twelve were rated by their teachers on the PF-nonPF measure. Since it is difficult at the high-school level to find students who follow the same program and at the same time to get a representative sample, it was decided to increase the total sample and to get test-retest data on the total sample, and to use a subsample restricted to a private school and urban schools where a combination of regular and student teachers was available for an inter-rater reliability check.

Teacher sample. The raters were twenty-two teachers from seven subject matter areas; twenty-seven class groups were included. The subject matter areas were: English, Social Studies, Science, Mathematics, Modern Languages, Secretarial Studies, and Shop.

Lists of teachers who had volunteered for the study were acquired, and the selection of those to be involved was based on the subject matter area of the teacher and the characteristics of the students. As representative a group of students as possible was chosen, in terms of age, sex, grade, and achievement level. In one sample--the largest--a modified randomization was used in the selection.

Class rosters were obtained from each participating teacher, and all class members were originally included in the sample. Those students with whom the teacher felt too little acquainted or those who were chronically absent were subsequently dropped from the sample.

Student sample. The final sample consisted of 610 JHS and HS students from seven New York City schools and two suburban schools. Originally, data were collected for 643 subjects, but for 33 subjects of one class data were incomplete and those subjects were only used for one special analysis of variance. Data for them were not included in the frequency distributions. The subjects were drawn from grades 9 through 12, and were between 13.1 and 19.3 years of age. Since no IQ scores were available, classes at various homogeneous achievement levels, as well as several heterogeneous-achievement groups, were selected to form a representative sample of high school students. The range of homogeneous class achievement levels was from "below average" to "above average." The distributions of subjects by age and sex, by grade, sex, and school location, and by level of achievement are given in Tables 1, 2, and 3.

TABLE 1
Distribution of Subjects by Age and Sex

Age in Years	No. of Boys	No. of Girls	Total no. Boys/Girls
13.1-14.6	55	62	117
14.7-15.6	70	34	104
15.7-16.6	75	52	127
16.7-17.6	87	71	158
17.7-19.3	51	53	104
TOTAL	338	272	610

TABLE 2
Distribution of Subjects by -- Grade and Type of School (N=610)

Grade	No. of Boys	No. of Girls	Total no. B/G	Type of Schools	
				Urban	Suburban
9	78	84	162	2	1
10	69	35	104	3	2
11	114	81	195	3	2
12	77	72	149	3	2

TABLE 3
Distribution of Subjects by Achievement Level

Achievement Level	No. of Subjects
Above average achievement	130
Average achievement	124
Below average achievement	154
Heterogeneous groupings	235
TOTAL	643(a)

Note.--Achievement level was determined by the Ss class designation.

(a) Data for 33 of the Ss in one class were incomplete and not included in subsequent distributions, thus leaving an N of 610.

The raters met with the investigator for two briefing sessions. Two considerations guided the briefing procedures of the teachers who acted as raters in the standardization of the instrument. One was that the teachers had an adequate understanding of the behavior to be rated, and the other, a sufficient acquaintance with the students.

At least six weeks after school had started briefing sessions of about one hour were held with the regular teachers at each of the cooperating schools and with the student teachers in the investigator's office. During these sessions, the rating instructions, a copy of the rating scale, as well as a sample of the trait rating sheet were distributed to the teachers. The rating instructions were first read out in the standard form so as to provide a common frame of reference. Then the scales were discussed one by one and questions were invited. The teachers were then asked to observe the behavior of their students in the classes to be rated for the next ten days to two weeks, after which they received the class rosters set up separately on the trait rating sheets for each of the ten subscales of PF-nonPF and the two ringer questions about achievement orientation and physical attractiveness. Special emphasis was placed on the instruction that in the rating they were to compare the students with one another as well as to keep in mind a general standard for these traits in adolescents in the high-school setting. This emphasis, of course, was meant to reduce the halo effect.

A second briefing session of about thirty minutes was held before the retest which was given after an interval of from four to six weeks. One additional feature in the second briefing session was the inclusion of a ranking sheet, on which the class was to be assessed along the teacher's global perception of PF-nonPF before the second rating was carried out. The ranking was added as an internal validity check. The above procedure was followed for all the teachers involved in the test-retest ratings.

A special subsample of nine teachers was involved in establishing interrater reliability on 158 subjects. Two school systems were involved, one public and one private; approximately half of the students came from each. Grades 9, 10, 11, and 12 were represented, and six sets of double ratings were obtained (that is, six sets of two teachers rated the same group of students so that their ratings could be compared). Of the nine teachers, three were student teachers and six were regular teachers. Three double ratings were obtained from the cooperating teacher and student teacher who worked with the children at the same hour of the day and in the same subject matter area. Three other double ratings were obtained from regular teachers who worked with the same groups of students but at different hours of the day and in different subject matter areas. Altogether, teachers in four different subject matter areas were involved. An attempt was made to hold briefing sessions on a more intensified scale with these raters. In two or three cases it was also possible to have a third session before the actual rating was done in order to clarify any points that might have come up in sample ratings they had been asked to do. However, in the cases where student teacher and regular teacher were paired, the briefing of the regular teacher was done "at second hand," i.e., by written instructions and by word of mouth through the student teacher. Double ratings for all groups except one were obtained at the time of the test.

Statistical Procedures

Since the primary goal of the study was the construction and standardization of a measure to assess playfulness and nonplayfulness in adolescents as observed in the high-school classroom, the statistical approach was aimed at defining playfulness and nonplayfulness by consensual agreement of the experts and then, on the basis of considering the emerging trait or traits continuous, using correlational analysis to determine the relationships among the separate PF-nonPF dimensions as well as those of age, grade and sex.

The dimensionality of PF-nonPF was explored by a principal components factor analysis, and was then, by varimax rotation reduced to four factors. Since test-retest correlations were sufficiently high (in the .70's), it was decided to use only the test data for the PF-nonPF ratings, and add to them the questions on achievement orientation and physical attractiveness as well as age and sex of student, grade level of student and sex of teacher which constituted a 16 x 16 matrix.

Some mention should be made here of the nonplayfulness end of the scale. Since this behavior crystallized as a complementary dimension and was built into the instrument in that manner, there could only be indirect evidence on the clusterings from individual profiles and from the frequency distributions of percentage scores.

The newness of the concepts involved, both in the formulation of the behavior by teachers and the investigator, as well as in rating the students later by their teachers, made it necessary to proceed with caution and to include validity and reliability checks at the various phases of the construction and standardization of the instrument.

The major correlation matrix for 610 Ss was a 35 x 35 table and consisted of the 10 PF-nonPF scores on test, the 2 ringer scores, sums of odd-numbered and even-numbered scales and total scale score, the retest scores on these 15 variables, age and sex of student, grade level, normalized ranking of PF-nonPF by teacher, and sex of teacher.

Reliabilities could, therefore, be read off directly from the matrix for each item with total score and one aspect of item reliability could thus be established. Of course, a certain amount of spuriousness enters into these coefficients and some caution must be exercised in their interpretation.

The reliability of the test itself was estimated by the Kuder-Richardson formula based on item statistics. Another check on reliability was obtained by correlating the quantity (A) Scales with the quality (B) Scales on the basis that frequency and degree of the same behavior should be equivalent.

A stability coefficient was also available from the re-test data, both for total test score, and by item (component traits) which provided a more conservative estimate of item reliability.

A separate matrix was run for 158 Ss on which ratings from two teachers were available. These data provided a measure of interrater reliability.

Validity data on the instrument consisted of the logical validity based on the content analysis and the concurrent validity obtained through the correlations of test, retest, as well as separate items with the rankings by the teachers of each subject on total PF-nonPF as conceived by the respective teacher.

Secondary considerations in the planning of the analysis of the data were the influence of age and sex of student, sex of teacher and subject matter area as well as the type of school system, namely whether urban or suburban.

Sex and age variables were fed into the major correlational analyses.

The influence of subject matter area and urban or suburban setting was explored by one-way analyses of variance.

As in all assessments by ratings, some thought had to be given to a possible halo effect. The most immediate clue available was again directly from the major correlation matrix, namely the coefficients of Scales VI and VII with total scale score. Another approach tried was to assess whether length of exposure to student and subject matter area, and/or both had any effect on ratings. Two matched groups were compared by a two-way analysis of variance in order to get additional data on halo effect.

III - RESULTS AND CONCLUSIONS

The results and conclusions deal with (1) the nature of playfulness-nonplayfulness in adolescents (2) the measurability of the behavior and (3) the influence of secondary variables.

The Nature of Playfulness and Nonplayfulness

Unidimensionality versus Multidimensionality of Playfulness-Nonplayfulness

The same question that was asked in the study with kindergarten children, namely, whether or not playfulness-nonplayfulness was a unitary trait, had to be explored at the adolescent level.

Because of intercorrelations in the 80's between A (quantity scales) and B (quality) scales at the kindergarten level, it was possible to combine the correlations and work with the five dimensions of physical, social and cognitive spontaneity, manifest joy and sense of humor as composite traits.

The pattern at the adolescent level was markedly different. Inspection of the major matrix (see Appendix C) shows intercorrelations among individual "A" and "B" scales ranging from a high of .53 (between IVA and IVB) to a low of -.13 between VA and VB, the mean value among all five scales being .40. Translated into behavioral terms, this suggests that the quantitative assessment of the social playful dimension of "group-oriented vs. self-oriented" showed a fairly high relationship to the qualitative ingredient of being "friendly vs. rejecting," but that if a student was rated, in cognitive playfulness, "intellectually alive," there was a slight but hardly significant tendency for him to be rated "conscientious," a nonplayful quality. When the relationships between the quantity and the quality of each dimension are compared to correlations among "A" scales or all "quantity scales" we find coefficients as high as .71 between scales IA and IVA, while the lowest is between IA and IIA, a correlation of .16. Looking now at the pattern of the "B" scales, we find that the highest r, .60, is between IIB and IVB, and the lowest, -.02, is registered for IB and VB. In view of these intercorrelations, no pooling of quantity and quality dimensions was carried out. Instead, the individual correlations were made the basis of the factor analysis. Since test-retest correlation was .82 for total test, and item test-retest correlations had a mean of .64, it was decided to use only test scores, i.e., the 10 ratings on PF-nonPF and the two "ringer" ratings on achievement-orientation and physical attractiveness as well as age of student, grade, sex of student and sex of teacher for the factor analysis.

The results of the principal components factor analysis and the subsequent varimax rotation to best fit may be examined for further clues to dimensionality. Four factors emerged as shown in Table 4.

Since the standardization of the instrument called for a cross-sectional, heterogeneous sample, certain cautions needed to be applied in the interpretation of loadings. The cut-off point for meaningful labelling was, therefore, set at .60. On this basis, distinct patterns can be seen to emerge in the first two factors. One factor is made up of physical mobility-physical rigidity, spontaneous joy-tenseness, humor-lack of humor, group orientation vs. self-orientation, friendliness-rejection, play-conscientiousness. The other factor consists of physical alertness (energy)-physical apathy, enthusiasm-discouragement, intellectual curiosity-intellectual stagnation, and the "ringer" question assessing ambition (achievement-orientation). Age and grade of student, and sex of student and sex of teacher emerged as two separate factors, unrelated to each other and the first two. The data from the factor analysis further support the decision not to pool A and B scales, since the first factor consists of Scales IA, IIB, IIIA, IVA, IVB, and VB, while the second factor is made up of Scales IB, IIA and VA. The two clusters of items suggest that two different kinds of playfulness-nonplayfulness are observable at the adolescent level, and that the dimensions making up one type are not major components of the other. These two types of PF-nonPF have been characterized as "social-emotional" PF-nonPF, a broad, situation-spanning behavior characteristic, and "academic" PF-nonPF, which appears more readily in the school situation (and is therefore narrower or situation-specific). As observed by teachers in the classroom setting, one could conceivably see academic playfulness as the teacher-approved type of playfulness which is constructive to the learning climate. Social-emotional playfulness-nonplayfulness would, by the same criterion, be held disruptive of the learning process.

That age and grade of student and sex of student and teacher have no appreciable effect on playfulness-nonplayfulness ratings was already intimated by the correlational analyses. The manner in which they crystallized as separate factors further confirmed the findings of minimal relationships between these variables and playfulness-nonplayfulness ratings.

Some mention should be made of the scales that showed lower than .60 loadings, yet were above the .30 loadings that are considered a level

TABLE 4

Rotated Factor Matrix for PF-nonPF and Ringer Scales,
CA, Sex of Student, Grade and Sex of Teacher
(N=610)

		Pattern loadings from Varimax rotations to best fit				Communality
	Rating	A	B	C	D	h^2
IA	Physical mobility-Physical rigidity	851	-050	-059	-023	732
IB	Physical alertness-Physical apathy	412	687	-057	-087	653
IIA	Enthusiasm-Discouragement	203	828	-025	005	727
IIB	Spontaneous joy-Tenseness	735	451	006	-023	746
IIIA	Humor-Lack of humor	803	278	021	-041	724
IIIB	Friendly wit-Hostile wit	502	413	102	-077	439
IVA	Group orientation-Self orientation	830	124	-071	077	714
IVB	Friendliness-Rejection	601	448	119	061	581
VA	Intellectual curiosity-Stagnation	213	810	-043	-068	707
VB	Play-Conscientiousness	669	-484	-105	-059	697
VI	Achievement orientation-Indifference	-260	825	-064	037	754
VII	Attractiveness-Homeliness	198	349	142	155	206
	CA	-035	-041	951	-019	908
	Sex of student	-098	-012	080	803	661
	Grade level	-014	016	948	-069	903
	Sex of teacher	066	009	-161	737	574

Note.--Decimal points have been omitted.

of practical significance in the interpretation of factors. Most conspicuously among these are the loadings of Scale IIIB, which aims at assessing the part of wit and subtlety in the adolescent's sense of humor as shown in the classroom. There is an almost equal distribution of saturation between social-emotional and academic playfulness, and it might indicate that perhaps two different traits have been combined--one, the teasing and off-color remarks, and the other, the more intellectual of punning and finding analogies. In the original content analysis of the Teacher Questionnaire the dimension "sense of humor (wit)" was seen as a combination of the other dimensions (social, cognitive, emotional and physical playfulness). The equal loadings of item IIIB on both PF factors suggests that such a combination was indeed being tapped. In a provisional formulation of PF as a self-rating instrument, the teasing and intellectual punning traits had been separated, and a correlation of .50 between them was found. It might be worthwhile to consider such a dichotomy once more.

Nonplayfulness. Because of the conceptual implications, the nonPF end of the scale was examined for further clues to the dimensionality of the trait. The question raised and to be answered is the nature of nonPF--complementary to PF or co-existant with PF.

The nonplayful ends of the scales were used by the teachers only slightly less frequently than were the playful ends. Over 20% of the students were rated "1" or "2", that is, primarily nonplayful, on each of the following scales: IA, IB, IIA, IIB, IVA, VA, and VB. For the remaining scales, the percentages rated "1" or "2" were as follows:

IIIA, 18.3% rated primarily humorless; IIIB, 10.2% rated primarily hostile or lacking in use of wit in class; and IVB, 11.2% rated primarily rejecting of peers. On scale VB, "erratic-conscientious," twice as many students were rated 1 or 2, conscientious, as were rated 4 or 5, erratic.

There were, nevertheless, no significant negative correlations among any of the PF-nonPF scales, indicating that nonplayfulness as assessed by these scales was not inversely related to the playfulness ends of the dimensions. In the rotated factor loadings, the only large negative loading was -.48 for scale VB, erratic-conscientious, on Factor 2, the academic PF factor. This loading suggests that the child who receives high PF ratings on the other items loaded on the factor, is likely to receive a low PF, or "conscientious" rating on item VB.

The conclusion to be drawn from the present data is that nonplayfulness does not exist as a separate entity identifiable through the pattern of intercorrelations.

In order to check on possible trends, the profiles of forty nonplayful students ranked as the lowest three by their teachers in fourteen different class-groups were examined. When the ratings of these students were tabulated, thirty-three of them showed a clustering on the nonPF end of the scale. However, it is the seven students whose profiles were uneven that may point to a less cleancut division between PF-nonPF behavior. Translated into percentage, it would mean that eighteen percent showed an overlapping of traits, a finding that calls for further investigation.

Measurability of the Behavior

Norms

The total student population consisted of 610 children, 338 boys and 272 girls, ranging in age from 13.1 to 19.3 years and attending grades nine through twelve. They were drawn from 7 urban and 2 suburban schools in New York City and surrounding areas. The percentage of the total group given each rating on each separate test item was computed. Test scores only were used, as test-retest correlations were uniformly high. The data constitute norms, to which other groups may be compared. The frequency distributions are presented in Table 5 below; the means

TABLE 5
Percentage of Subjects Given
Each Rating on the PF-nonPF and Ringer Scales
(N=610)

	Ratings				
Subscales	1	2	3	4	5
IA	5.25	19.18	52.46	16.39	6.72
IB	2.46	22.31	49.02	20.66	6.56
IIA	2.46	18.85	46.23	23.11	9.34
IIB	3.93	18.20	46.39	18.20	13.28
IIIA	2.62	15.74	55.08	17.54	9.02
IIIB	1.64	8.69	64.43	18.85	6.39
IVA	6.23	20.82	45.90	19.02	8.03
IVB	1.64	9.51	53.11	23.93	11.80
VA	4.43	20.33	45.08	20.98	9.18
VB	9.51	24.92	48.36	13.28	3.93
VI	1.80	16.72	49.34	23.11	9.02
VII	2.95	12.13	55.41	21.31	8.20

Note.--N=610. A rating of 1 denotes the nonPF end of the scale; a rating of 5 denotes the PF end of the scale.

and standard deviations in Table 6. For analyses by age, sex, or grade, trends discussed in those respective sections should be considered as additions to and modifications of these data.

TABLE 6
Means and Standard Deviations on PF-nonPF and
Ringer Scales for Test and Retest
(N=610)

	Subscales	Test		Retest	
		Mean	SD	Mean	SD
IA	Physically on the move: Physically rigid	3.00	.91	3.02	.99
IB	Physically alert: Physically apathetic	3.08	.88	3.13	.94
IIA	Enthusiastic: Discouraged	3.18	.93	3.18	.94
IIB	Relaxed (Spontaneous): Tense (Constricted)	3.19	1.01	3.20	1.01
IIIA	Fun-Loving: Humorless	3.14	.88	3.14	.94
IIIB	Accepting in wit: Hostile in wit	3.20	.75	3.18	.84
IVA	Group-Oriented: Self-Oriented	3.02	.98	3.02	1.01
IVB	Friendly: Rejecting	3.35	.87	3.34	.90
VA	Intellectually alive: Intellectually stagnant	3.10	.97	3.10	1.02
VB	Erratic: Conscientious	2.77	.93	2.86	.94
VI	Ambitious: Indifferent	3.21	.89	3.21	.88
VII	Beautiful (Handsome): Plain (Unattractive)	3.20	.86	3.25	.87

All of the distributions are unimodal and in all but one case between 45% and 55% of the students are rated "3." Scale IIIB has less variability than the other scales; over 64% of the students are rated "3." The distributions are almost exactly balanced for items IA, IB, and IVA. There are slight skews toward the PF rating for items IIIA and VA and toward the nonPF end for scale VB; in each of these the difference is less than ten percentage points between 1-2 and 4-5 ratings. In scales IIA, IIB, IIIB, IVB, VI, and VII there is a more marked skew toward the PF ends of the continua and the positive ends of the "ringer" questions. These last results indicate that teachers rate more of their students as enthusiastic, relaxed, accepting in wit, friendly, achievement-oriented, and attractive than their opposites. Table 6 shows that the mean of the ratings on all scales except VB is "3" or above. The latter result indicates that teachers rate a majority of students as "conscientious" rather than "erratic" in studies. The data on sex differences indicate that girls, especially, are more likely to be rated "conscientious."

The striking similarity between test and retest means indicates that, although teachers' ratings of individual students varied somewhat over time, their overall patterns of use of the scales did not change. The preponderance of "3" ratings on item IIIB suggests that the use of wit does not occur often enough in the high-school classroom for the teachers to make meaningful discriminations among the students.

Validity

Since the behavior to be identified was a new concept, the problem of validity was crucial. Two approaches to validation were explored.

Rational (logical) validity. The rational basis for the formulation of the instrument was the content analysis of the 115 open-ended questionnaires completed by JHS and HS teachers.

As had been indicated in the discussion under "Construction of PF-nonPF Scale," the decision to use the teachers' criteria was dictated by the desire to establish face validity for the behavior as seen in the classroom. The criteria used by the investigator and those suggested by fellow psychologists were additional checks on consensual validity.

A reading and first sampling of 36 questionnaires showed a distinct dichotomy in valence of traits and incidents describing both playfulness and nonplayfulness. At one point, it was thought that a score or rating could conceivably be based on summing positive and negative mentions but coder agreement was not achieved to a satisfactory degree. It was therefore decided to combine traits regardless of valence and put the emphasis on logical consistency of behavior. A frequency count of the behavioral correlates mentioned by the teachers produced 294 mentions for playfulness and 260 for nonplayfulness. A preliminary sort of these correlates of playfulness-nonplayfulness as observed in the classroom produced ten categories for playfulness and 21 for nonplayfulness. The investigator and the graduate assistant then aimed at establishing logically consistent profiles within the teachers' formulations and were able to reduce the 21 categories of nonplayfulness to 10. Inspection of these categories suggested a complementary trend, and these ten profiles for playfulness and nonplayfulness were matched and formed the basis of the ten playfulness-nonplayfulness scales.

At this point it was possible to compute percentage agreement by the teachers on the traits making up the profiles. The agreement ranges from a high of fifteen percent on the fun-loving aspect of playfulness to a low of two percent for friendly on the playfulness end of the scale, and from a high of twenty-three percent on self-oriented to slightly over one percent on physically rigid on nonplayfulness. The percentages for the remaining eight scale traits are given in Table 7.

TABLE 7
Percentage of Teachers Mentioning Traits Used in the Formulation of the PF-nonPF scale (N=115)

	Playfulness (PF)		Nonplayfulness (nonPF)	
Number	Name of Trait	% Mentioning Trait	Name of Trait	% Mentioning Trait
IA	Physically on the move	14.69	Physically rigid	1.15
IB	Physically alert	2.04	Physically apathetic	1.93
IIA	Enthusiastic	6.12	Discouraged	11.53
IIB	Relaxed	9.52	Tense	12.31
IIIA	Fun-loving	15.30	Humorless	10.77
IIIB	Accepting in wit	13.60	Hostile in wit	5.38
IVA	Group-oriented	12.24	Self-oriented	22.70
IVB	Friendly	8.84	Rejecting	1.93
VA	Intellectually alive	12.92	Intellectually stagnant	16.92
VB	Erratic	4.42	Conscientious	15.35

The decision to accept what might at first glance appear to be low percentage agreement reflects the great diversity of the traits mentioned under the concepts of playfulness-nonplayfulness. In order to do justice and capture some of the nuances, no definite cut-off point was set.

Concurrent Validity

To determine concurrent validity of the test, a comparison was made between the normalized rankings the teachers assigned their students on overall playfulness and the students' scores on the playfulness measure. Although the teachers' rankings were obtained after the first test was administered, it can be considered an adequate criterion for the following reasons. First, the rankings were obtained at least four weeks after the first ratings, when the teachers' original scorings were not fresh in their memory. At the same time, having been asked to evaluate the playfulness of the students once on specific subscales, it can be expected that they were sensitized to behavior that might be termed "playful" and thus able to evaluate their students more discriminatingly. Further, it is assumed that the tasks of ranking all the students on a global "playful" dimension on the one hand, and rating each individual student on each of ten subscale continua on the other, are sufficiently different to minimize the biasing effect of one on the other. This interpretation is supported by the correlations between the rankings and the halo items, which were as follows:

Rank and achievement-orientation (test): .00
Rank and achievement-orientation (retest): -.07
Rank and physical attractiveness (test): .21
Rank and physical attractiveness (retest): .19

In contrast, validity coefficients as assessed by the correlations between teachers' rankings and total scale and subscale ratings on PF-nonPF components are uniformly high as shown in Table 8.

TABLE 8
Correlations Between Ss' Rank on Playfulness (Global) and Ss' Score on PF-nonPF Subscales for Test and Retest
(N=610)

		Correlations with Rank	
	Subscales	Test	Retest
Total score		.69	.76
IA	Physically on the move: Physically rigid	.58	.72
IB	Physically alert: Physically apathetic	.44	.62
IIA	Enthusiastic: Discouraged	.32	.35
IIB	Relaxed (Spontaneous): Tense (Constricted)	.59	.73
IIIA	Fun-loving: Humorless	.63	.71
IIIB	Accepting in wit: Hostile in wit	.43	.51
IVA	Group-oriented: Self-oriented	.58	.68
IVB	Friendly: Rejecting	.44	.54
VA	Intellectually alive: Intellectually stagnant	.36	.35
VB	Erratic: Conscientious	.33	.42

The generally higher correlations of the ranking scores with retest scores may indicate a carry-over from the rankings.

Reliability

Both total test scores, split-half scores and item scores were analyzed for reliability. Data are, therefore, available for internal consistency, stability over time, and item consistency and stability. A special section deals with interrater reliability.

Reliability coefficients of internal consistency were obtained using the Kuder-Richardson formula based on test variance. These were .87 on test and .90 on retest.

Since quantity and quality scales were considered equivalent, reliability coefficients by the split-half technique between A and B Scales were .84 on test and .86 on retest. Corrected for length by the Spearman-Brown formula, the split-half reliabilities are .91 on test and .92 on retest.

A coefficient of stability over time was obtained for the total test scores with total retest scores in the amount of .82. Test-retest coefficients for A-Scales was .80 and for B-Scales .75. Test-retest reliabilities for each scale (item stability) are shown in Table 9.

The correlations of each item score with total test score were also inspected for reliability data. Although these correlations are inflated because the subscale score is a component of the total score, the correlations further support the internal consistency of the measure. Item reliabilities on test ranged from .36 to .86, with a mean of .68. On retest, the range was from .34 to .89, with a mean of .74.

Mention should also be made of the reliability of the "halo" items, namely, Scales VI and VII. Test-retest stability for Scale VI, measuring achievement orientation, was .71, and for Scale VII, measuring physical attractiveness, it was .78.

TABLE 9
Item Reliability Coefficients Between
PF-nonPF Subscale Score on Test
and PF-nonPF Subscale Score on Retest (N=610)

Scale		r
IA	Physical mobility	.68
IB	Physical alertness	.62
IIA	Enthusiasm	.64
IIB	Spontaneous joy	.68
IIIA	Humor	.67
IIIB	Wit	.55
IVA	Self-Other orientation	.69
IVB	Friendliness	.59
VA	Intellectual curiosity	.68
VB	Erratic behavior	.60
Mean r		.64

Interrater Reliability. For a subsample of 158 students, separate ratings by two teachers were obtained to investigate the interrater reliability of the instrument. Two school systems were involved, one public and one private; approximately half of the students came from each. Grades 9, 10, 11, and 12 were represented, and six sets of double ratings were obtained, 3 from the private and 3 from public schools. Altogether, 9 teachers were involved, of whom 3 were student teachers and 6 were regular teachers. Three double ratings were thus obtained from the cooperating teacher and student teacher who worked with the children at the same hour of the day and in the same subject matter area. Three other double ratings were obtained from regular teachers who worked with the same groups of students but at different hours of the day and in different subject matter areas. Altogether, teachers in four different subject matter areas were involved (see Table 10).

TABLE 10
Distribution of Ss in the Interrater Reliability Design by Subject Matter Area, Teacher Status and Order of Rating (N=158)

Status of Teachers			Subject Matter Area	
First Rating	Double Rating	N of Ss	First Rating	Double Rating
Regular	Regular	24	English	Science
Regular	Regular	19	English	Mod. Lang[a]
Student	Regular	27	English	English
Regular	Regular	29	Modern Languages	Science
Regular	Student	31	Social Studies	Social Studies
Student	Regular	28	Social Studies	Social Studies

(a) Modern languages taught were Hebrew and Spanish.

This mixed sample was the only one from which double ratings could be obtained, and it was felt that interrater reliabilities from the sample might be minimal. The student population in the private school was much more homogeneous than was the entire sample of students, both in socio-economic status of the home and in level of academic achievement; thus homogeneity of half of the population was also expected to minimize interrater reliability.

The correlations between ratings by two teachers of the 158 students are shown in Table 11.

TABLE 11
Interrater Reliability Coefficients on PF-nonPF
and Ringer Scales (N=158)

	Scale	r
IA	Physically on the move:Physically rigid	.30
IB	Physically alert:Physically apathetic	.39
IIA	Enthusiastic:Discouraged	.34
IIB	Relaxed(Spontaneous):Tense(Constricted)	.52
IIIA	Fun-Loving:Humorless	.47
IIIB	Accepting in wit:Hostile in wit	.23
IVA	Group-Oriented:Self-Oriented	.47
IVB	Friendly:Rejecting	.28
VA	Intellectually alive:Intellectually stagnant	.46
VB	Erratic:Conscientious	.29
VI	Ambitious:Indifferent	.54
VII	Beautiful(Handsome):Plain(Unattractive)	.31

It is interesting to note that the highest interrater correlations occur for the dimensions "relaxed-tense," "fun loving-humorless," "group vs self-oriented," "intellectually alive-intellectually stagnant," and "ambitious-indifferent," the last a "ringer" item. It may be conjectured that these qualities in students are the most easily observable for teachers, and that on the one hand the student's intellectual liveliness and ambition are of special concern to teachers, who try to encourage these qualities in the students. On the other, the more disruptive elements of the fun-loving and chatty, gregarious youngster seems to stand out equally clearly to the teachers.

In two instances, a pair of teachers working with the same class had also ranked their students on the overall dimension of playfulness. Interrater reliabilities were determined using a Spearman-rank correlation. One pair were regular teachers, both male, who instructed the group in different subject matter areas; the correlation between their rankings of the 19 students was .34, which is not significantly different from zero at the .05 level. The other pair were a regular teacher and her student teacher, both female, who saw the students during the same class period and instructed them in the same subject matter area. The correlation between their rankings of the 31 students in their class was .27, again not significantly different from zero.

Therefore, satisfactory interrater reliability cannot be claimed for the present data, either for specific ratings on PF-nonPF subscales, or for an overall conception of PF-nonPF. In part, it may be assumed that the many differences among both student populations and raters contributed to the unreliability of the data. It may be, also, that students' behavior in various classes and in response to different teacher personalities led to consistent differences in PF-nonPF ratings of the high school student. While satisfactory reliability might be obtained eventually on some of the more "overt" scales, those scales which rely on more "covert" characteristics may remain unreliable when comparisons are made among teachers.

It cannot be determined from the present data, which of the many factors may have contributed to the low reliability, but the comparison of rankings of overall PF-nonPF between a pair of regular teachers and a cooperating teacher-student teacher pair suggests that student teachers and regular teachers show the same inconsistencies.

Halo

Directly related to both validity and reliability is the question of bias in ratings, which from student to student may affect validity, and in ratings of the same student, be a "generosity" error affecting reliability.

To assess any carry-over that might occur from an overall general perception of the student on playfulness-nonplayfulness, to what were considered the "ringer" questions on "ambitious-indifferent" behavior and the trait of "beautiful (handsome)-plain," the correlations between these questions and total test coefficients on test and retest were inspected and are .17 and .12 for scale VI and .26 and .28 for scale VII, allowing the deduction that there was only limited carry-over.

Another indication of the halo effect or the lack of it can be found in the data on concurrent validity. Teacher-ranking of playfulness-nonplayfulness as a global concept showed with total test score coefficients of .69 and .76 on test and retest, respectively The range of item validity coefficients was from .32 to .63 on test, and from .35 to .72 on retest, with respective means of .47 and .56. Coefficients on test and retest between achievement orientation and teacher-rank was zero, and between physical attractiveness and teacher-rank .20.

The conclusion may be drawn that the playfulness-nonplayfulness conceptualization of the teachers had very little, if any, effect on their ratings of the "ringer" items.

In the original design, comparing high-and low-exposure of teachers to students was also aimed partly at getting information on the halo effect on PF-nonPF ratings. The two-way analysis of variance shows that length of exposure of students to the teachers did not affect teachers' ratings on Scales IB, IIB, IIIA, and IIIB. Behaviorally this means that "physically alert-physically apathetic," "relaxed-tense," "fun-loving-humorless," "accepting in wit-hostile in wit," are not affected by length of acquaintance with the student.

It might be worth noting that except for the scale on physical alertness-physical apathy, the dimensions which were not influenced by exposure are all more heavily loaded on the emotional-social factor of PF-nonPF, which is also interpreted as situation-spanning. Conversely, one might draw the conclusion that academic playfulness is more contaminated by the teacher's image of the student as a learner and might therefore be more susceptible to the variable of exposure, i.e., knowing him well or hardly at all.

Results of the factor analysis are another source for information on halo effect. The loadings of Scales VI and VII on Factor 1 (social-emotional PF) are -.26 and .20 respectively, and on Factor 2 (academic PF) .82 and .35. Since any factor loading below .30 need not be considered significant, it is clear that there is no carry-over from social-emotional PF-nonPF to the ringer questions. Yet, the picture changes dramatically when we examine the relationship of academic PF to the ringer questions. Scale VI, "ambitious-indifferent" is highly saturated with that particular factor, and there is even a slight carry-over to Scale VII, beautiful (handsome)-plain (unattractive).

We can, therefore, infer that the very nature and the very distinct difference between the two PF factors is reflected in the differential halo effect. Teachers are more likely to have an overall favorable impression of the academically conspicuous boy or girl. In addition, some behavioral correlates of ambition as expressed in Scale VII can be found in the profile of the physically alert, enthusiastic and curious student.

Another kind of "halo" sometimes occurs in a "generosity error," that is, raters use predominantly what they consider to be the "favorable" end of a rating continuum. From the frequency distributions, it can be seen that this did occur on some scales, including the "halo" items, but that the most frequently used rating on each item was "3," and that the distributions were only slightly skewed if at all.

Influence of Secondary Variables:
Age, Grade, Sex, Subject Matter Area, Type of School

As already mentioned, the present investigation wanted to take into account the influence of some developmental and environmental variables on PF-nonPF behavior in HS students.

Age

The underlying assumption about the influence of age was that children of different ages might be given different playfulness ratings by their teachers. Accordingly, age was one variable in the correlation matrix, and the data were cast in frequency distributions for each item according to five age groupings (see Table 1).

When age was correlated separately with each of the test and retest items and total playfulness-nonplayfulness scores, all correlations were approximately zero. The range of correlations with the test items was -.10 to +.04, with the retest items, -.12 to -.01. Correlations of age with total playfulness scores were -.06 and -.10 for test and retest respectively. Correlations between age and "ringer" items were also close to zero on both administrations of the test. (This would be expected, from the high test-retest reliabilities of the ringer items.)

Inspection of the frequency distributions indicated that, although distributions varied in shape from item to item (see Norms), there were no differences among age groups in the shapes of the distributions on any playfulness-nonplayfulness item.

It is therefore concluded that, in this population, teachers did not rate different age groups differently on playfulness-nonplayfulness. It might be inferred from this finding that throughout adolescence the same kinds of playful behavior are engaged in, at least so far as is evident to teachers. Another interpretation is that the teachers rated their classes using only the particular age group in question as a reference group, and thus rated each child relative only to that group. If the latter occurred, it would be expected that approximately normal distributions on each item would occur in each age group, and no comparison could be made across groups for relative frequencies in the occurrence of kinds of playful-nonplayful behavior. There is also the possibility that the classroom as a setting might act as a leveller of behavior.

One exception to this trend was further explored. It was determined by inspection that significant age differences might have occurred in ratings on Scale VI, "ambitious-indifferent." A chi-square analysis of observed in contrast to expected frequencies was performed on a 5x3 table (5 age groups x 3 levels of ratings) and showed a departure from chance expectation significant at the .01 level. It is concluded that ratings on achievement orientation are not independent of age, but that younger students are more often rated "3" than are older students, and older students are more often rated toward one or the other end of the continuum than are younger students. It could be conjectured that in the latter years of HS this behavior becomes more salient.

The correlations between grade level and test and retest scores ranged from -.09 to .03. The correlations with total scores on test and retest were, respectively, -.01 and -.07. Thus, again, within-group variability was very wide and overshadowed any potential between-group differences.

Group trends appeared, however, on inspection of the frequency distribution by grade level. The sex differences associated with age appeared in the same patterns in the distributions by grade, although some differences were reduced, partially due to the recombination of the data into four rather than five groups.

As grade level increased, greater variability among scores occurred on two subscales: physically on the move-physically rigid, and accepting in wit-hostile in wit. In the upper grade levels, more students were rated self-oriented as opposed to group-oriented than in the lower grade levels. These were the only apparent trends in the data, suggesting that there are no major differences in PF-nonPF among adolescents according to their grade levels. This conclusion was supported in the factor analysis, where age and grade emerged as isolated factors with no loadings on the playfulness-nonplayfulness factors.

Sex

The sex of the student was correlated with each of the test items and with the total scores. The range of correlations, in which positive correlations indicate high PF-nonPF ratings for girls, and negative correlations indicate high PF-nonPF ratings for boys, was from -.14 to -.01 on the test and from -.13 to .04 on retest. Correlations of the ringer item ratings with sex of pupil were positive but none was greater than .06 and thus not significantly different from zero.

It was concluded from these correlations that boys and girls were not rated differently on overall playfulness or on any of the specific subscales.

Inspection of the frequency distributions, which grouped the students into five age groups and four grade levels, shows however some trends toward consistent sex differences.

Boys' ratings on being physically on the move had a wider variability than did girls'. Girls' ratings on physical alertness had a wider spread than did boys'; there was a slight tendency for teachers to rate more girls as physically rigid, although the percentages of ratings of high mobility were equivalent.

The spread of scores on the relaxed-tense dimension was wider for boys, and boys were generally given more 4 and 5 ratings than were girls. Grouped by grade level, boys were given proportionately more 4 and 5 ratings in each group, but the difference was greatest in the ninth grade.

On the dimension fun-loving - humorless, boys were given more playful and girls more nonplayful ratings in the two younger age groups, but these differences largely disappeared in the older groups. Boys were also given more playful ratings on the subscale accepting-hostile in wit, and the variability in ratings was wider for boys than girls. The same pattern was observed in scores on the group-oriented - self-oriented scale, especially for grades 11 and 12.

On the dimension intellectually alive-intellectually stagnant, there was again a wider spread among boys' ratings; the younger boys were rated "alive" more often than were the girls, but the latter difference was not observed among older children.

Girls at each age level were more often rated "conscientious" than were boys, and boys were more often rated "erratic," i.e., at the playful end of the continuum.

These data suggest that boys and girls do engage in different amounts and kinds of playful behavior, and that boys are seen by their teachers as being more playful than girls in several areas. The very low intercorrelations of the PF-nonPF scores with sex indicate that variability within each sex group is much wider than between-group variability; however, group data point to consistent sex-role trends.

Subject Matter Area

The students were studied in class groups which, altogether, included seven subject matter areas. A total of 22 teachers taught 27 class groups. The subject matter areas represented and the total number of students rated in each are shown in Table 12.

TABLE 12
Distribution of Students and Teachers by Subject Matter Area
(N=610)

Subject Matter Area	No. of Students	No. of Teachers
English	176	5
Social Studies	118	5
10th Yr. Mathematics	85	2
Science	98	4
Modern Languages(a)	51	2
Secretarial Studies	64	3
Shop	19	1

(a) The modern languages taught were Hebrew and Spanish.

The secretarial studies classes consisted primarily of girls, and the shop classes primarily of boys.

A one-way analysis of variance was performed on each subscale to ascertain whether consistent differences occurred among ratings by teachers in different subject matter areas. Except for Scales IB and IIA, differences among subject matter areas were significant at the .01 level on all PF-nonPF items. The difference was significant at the .025 level for item IB, and was not significant for item IIA.

It is concluded that PF-nonPF ratings are consistently different according to which area of the curriculum the teacher instructs the students in. In order to ascertain the direction of the differences, the mean ratings in each subject matter area for each PF-nonPF item were ranked. Although the means differed very little from one another, there was a consistent trend for highest ratings to be given in shop, followed closely by English and Modern Languages, and for lowest ratings to be given in Mathematics and Secretarial Studies. Since the Shop classes were made up exclusively of boys, and the Secretarial Studies classes of girls, the hypothesis that a sex difference exists, either in students' behavior or teachers' perceptions or both, such that boys are rated more playful and girls less so is supported.

The reason for the observed differences among subject matter areas might lie in either of two directions. It may have been that students acted differently in class in response to the subject matter area itself: perhaps some curriculum areas lend themselves differently than others to playfulness or nonplayfulness behavior. On the other hand it may have been that teachers' personalities were consistently different between subject matter areas, and that teachers and students were responding differentially to each other. It seems plausible that major personality differences which might influence perception and/or attitude toward playfulness might occur between, for instance, English teachers and Mathematics teachers. Significant differences between subject matter areas at at least the .05 level were obtained on the halo items, suggesting that there is an overall difference in perception of students between subject matter areas and that teacher personalities may be the base of the observed differences.

It should be noted that the data from the two-way analysis of variance are consistent with these. Although the subject matter area variable is not a significantly differentiating factor in all of the scales in which it did produce a significant difference in the one-way analysis, the magnitude of the F-ratios stand in the same relationship to each other in the two analyses. The failure of some to reach significance in the two-way analysis may be attributed to the smaller sample and thus much lower power of that test.

Type of School

Seven urban schools were represented by 369 students and two suburban schools by 241 students. A one-way analysis of variance performed to assess differences in PF-nonPF ratings which might occur between urban and suburban schools, showed that there were none, either on the PF-nonPF subscales or on the halo items. This suggests that the overall characteristics of schools, their teachers, and their students, do not vary along an urban-suburban continuum in characteristics relevant to playfulness-nonplayfulness.

IV - DISCUSSION

The discussion will focus on (1) the nature of PF-nonPF in adolescents; (2) the measurability of the behavior labeled PF-nonPF; and (3) the influence of secondary variables on the trait.

The Nature of PF-nonPF in Adolescents

Psychological Meaning of Two-Factor Composition

The most significant finding is the emergence of a two-factor syndrome in PF-nonPF in adolescents as observed in a classroom setting. The factor labeled "academic playfulness" with its high loading on the "ringer" question of achievement-orientation suggests that we might deal here with a situation-specific manifestation of playfulness. Looking more closely at the other trait loadings of physical energy (alertness), enthusiasm and intellectual curiosity one might even be tempted to call this cluster (behavior) the teacher-approved type of playfulness. In the light of the observations by Wallach and Kogan who saw their more creative pupils as showing a playful attitude toward knowledge, it will be interesting to watch the relationship of these traits to the divergent and convergent thinking measures in the planned follow-up study. At this point, two interpretations suggest themselves. One relates to the original conceptualization of PF-nonPF by the teachers. Getzels and Jackson (1962) among others found that teachers were predisposed to the intelligent, convergent thinker who would be the youngster more concerned with excellence in grades and tests. If this is so, then the contamination of Factor 2 of PF-nonPF with high-achievement orientation might support this bias by the teacher. The other interpretation is based on the repeated statements and findings by researchers (Thorndike 1963; Wallach and Kogan 1965; Denny _et al_ 1967) that there is an overlap between and among the abilities labeled as convergent and divergent. Evidence to this effect was also found in our study with kindergartners (Lieberman 1965) but the equivocal results there could be attributed, at least partially, to a lack of differentation of abilities in the young child.

Looking now at the factor labeled "social-emotional" PF-nonPF with its loadings on physical mobility, spontaneous joy, humor, sociability, and "play" we might conjecture that this is behavior that would be considered disruptive in a classroom setting and therefore not tolerated by the teacher, or perhaps more accurately, by a majority of teachers. Again the observations and findings in the literature (Wallach and Kogan 1965; Getzels and Jackson 1962) would support this interpretation. Further evidence for a somewhat ambivalent attitude on the part of teachers comes from the analysis of the answers to a sample of the open-ended T-questionnaire, which showed negative connotations in one-third of the behavior mentioned as playfulness in classroom behavior. This contrasts with only about one-fourth negative traits when applied to a teenager in general. The latter finding also supports our contention here that the social-emotional element in PF-nonPF might be situation-spanning, i.e., more likely to be manifested outside the classroom as well and considered less noxious by the teacher there.

Calling one factor situation-specific and the other situation-spanning suggests that the traits descriptive of PF-nonPF in a HS student need to be put to the empirical test outside the classroom. Additional data assessing PF-nonPF in adolescents in a leisure type setting are, therefore, being collected.

McReynolds (1964), in his work with an individual's preferred innovation rate (PIR) and in his analysis of fun, mentions also an individual's innovation style as a possible explanation for differences in the type of activities that afford him pleasure. He cites the ideational and motoric styles as two possible manifestations. In trying to identify and measure PF-nonPF in the classroom, our hope was to isolate some indices of "fun in learning" even in a structured situation. The two factors of PF-nonPF could therefore point to different approaches and, in line with Guilford's (1956) model of the intellect, might give a clue to creativity in different content areas, including the behavioral.

Another aspect of viewing the two-factor constellation of PF-nonPF in adolescents is to compare it to the behavior labeled "impulsivity" in the work of Kagan(1966), and Sutton-Smith and Rosenberg (1961). There is a definite need to sharpen and underline the difference of the traits subsumed under PF-nonPF and those listed under impulsivity. Very often not only the layman but also professionals working with children and adolescents confuse the "healthy" aspects of playfulness with some of the "maladaptive characteristics" of impulsivity. Wallach and Kogan (1965) took exception to Kagan's thesis that only the analytical mind could produce the higher type of thought and argued for a different look at the relational processes of thought. Some of the controversy about the interpretation of "relational" as it pertains to creative processes has, in the meantime, been resolved through a sharpening of the type of combinations covered by that label (Kagan 1967). Similarly, attention-getting devices such as calling out and waving one's hand in class have been found as distinguishing features of high creative girls (Wallach and Kogan 1965) and would point to the validity of the "physical energy" ingredient in academic playfulness. Another clue to a separation of disruptive from playful behavior is the finding that while 12% of the teachers agreed that calling-out and pushing and shoving constituted a trait in playfulness, this particular trait did not reach a significant loading of .60 on either Factor 1 or Factor 2.

Comparison of Behavior Correlates of PF-nonPF at the HS Adolescent and Kindergarten Levels.

Concern with trait continuity over time, and stability of a trait in an individual, is reflected in the work of Emmerich (1964) and Kagan and Moss (1962). While continuity over a time may be explored through cross-sectional design, stability of a trait in an individual requires, of course, a longitudinal study. In his analysis of the behavior of nursery school children, Emmerich demonstrated that observations during free play could be studied along a continuity-discontinuity dimension over time in relation to individual stability. Kagan and Moss used the term of "developmental transformation" in their findings of change in behavioral correlates, i.e. in male passivity, in the subjects of the Fels longitudinal study. By using a more encompassing concept of masculinity they claimed continuity despite phenotypic change.

The long-term design of PF-nonPF as a quality of play and of the player has as one of its aims a comparison of the behavioral indices of PF-nonPF at various age levels. At the present stage of the investigation, data are available from cross-sectional studies of HS students and kindergartners which allow a descriptive comparison between the component traits of PF-behavior at these two levels for possible clues to continuity over

(2)

time. Since no structured definition of nonplayfulness was used at the kindergarten level, the comparison must restrict itself to the playfulness dimensions.

The most striking alikeness occurs in the manifestations of playfulness across physical, social, emotional and cognitive activities. The differences are most evident in the negative connotation of playfulness behavior at the adolescent level as well as the two-factor structure in the HS classroom as opposed to the unitary dimension found at the K-level.

As to the first finding, the question may be legitimately raised whether the conceptualization across different areas of functioning was influenced by the investigator's earlier formulation at the K-level. In this connection, it should be remembered that the basis of the PF-nonPF measure at the A-level was the teachers' criteria and that the categories were independently checked by a graduate assistant who was not familiar with the K-scale. As one of the undergraduate students who collected data for playfulness indices in a leisure-type setting remarked, "It is a mode of approach to the environment."

A more rigorous comparison can be made from the factorial findings which might allow the drawing of more definite conclusions. At the K-level, a heavy saturation of all scales with the first centroid factor pointed to a unitary trait for playfulness in the kindergartners sampled. Orthogonal handrotations suggested that in addition to the factor of spontaneity in playfulness, there were two secondary factors, namely, brightness and maturation. The picture is very different at the adolescent level. The separation of physical alertness, enthusiasm and intellectual curiosity into one cluster, and physical mobility, spontaneous joy, humor and sociability into another, might be interpreted as more differentiated behavior and more situation-specific at an older age. The absence of a maturational factor at the adolescent level needs to be further tested. There is just a possibility that it may be an artifact of the raters' frame of reference, by ignoring an absolute standard and only using their class group as a yardstick. On the other hand, it might be a developmentally valid finding and point to crystallization of the trait in early adolescence and a stability over time. Only longitudinal data could give a clearcut answer to this hypothesis.

Some mention needs to be made of the trait "sense of humor" in kindergartners and "wit" in adolescents. While "glint-in-the-eye" behavior in young children was related to the teacher's estimation of intelligence, analogous behavior in the adolescent showed about equal loadings on the academic and social-emotional factors of playfulness, neither of which reached the cut-off point of .60 for meaningful psychological interpretation. Perhaps such behavior might either be too difficult to observe by the classroom teacher or too threatening for the classroom climate to be allowed free play.

Nonplayfulness

In positing a behavior dimension of playfulness, it was necessary to consider the possibility of traits that could be labeled "nonplayful." From the frequency distributions it becomes apparent that nonplayful traits exist in HS students to about the same extent as do playful dimensions. The question that is not answered by the data is whether a clustering exists at the nonplayfulness end similar to that suggested by the factorial analysis for the positive aspects, i.e., playful behavior. Since the teachers' conceptualization suggested a complementary trait rather than coexisting behavior and since no definite clustering of nonplayfulness other than the opposite of playfulness as formulated is available from the data, it seems justified to use the factorial composition of the nonplayfulness end of the scale as one clue. We can, therefore, try to speak of social-emotional nonplayfulness and academic nonplayfulness.

The nonplayful syndrome of physical rigidity, tenseness, lack of humor, and rejecting behavior and "conscientiousness" as opposed to a nonplayfulness consisting of physical apathy, discouragement, intellectual stagnation and indifference seems to outline two different individuals or as in our case, two different students in the HS classroom. Studies in rigidity of set in problem solving (Cunningham, 1965) may not have paid sufficient attention to the social-emotional vs. the intellectual ingredient in "Einstellung Rigidity." It would seem possible to have a highly intelligent, rigid youngster whose creative potential may be undercut by social-emotional factors. Conversely, no amount of flexibility will spark a youngster with low intelligence in the academic subjects. If as Wallach and Kogan maintain, there is a group of high-creative:low intelligence youngsters, then possibly their creativity needs to be directed to other "content areas" as indicated by Guilford's (1956) model. A follow-up study separating the playfulness and nonplayfulness attributes is designed to take this discussion out of the speculative realm and provide some statistical answers to the conjectures put forward here.

Another possibility would be to use the individual profiles of the forty cases at the extreme end of the nonplayful ranking and examine them further for a cluster analysis. However, the present data tabulation would make this into a too cumbersome procedure and it is planned to make provisions for such an analysis in a different approach to evaluate PF-nonPF by a trait check list rather than a rating scale.

At this point in the investigations, the question of what is nonplayfulness and how it can be assessed remains only partially answered.

Measurability of Behavior

As in the K-study, it was of paramount importance to test whether the behavioral correlates of PF-nonPF can be validly and reliably measured. It was also necessary to find out whether individual differences would be meaningfully recorded.

Norms

To the extent that the population sample was representative of a cross-section of urban and suburban HS students, the mean ratings can be referred to as norms. The most striking features are their uniformity in values and the fact that the mean is the midpoint of the scale for all scales but VB. Since the standard deviation is approximately 1.00 for all scales but IIIB, the satisfactory variability precludes the suspicion that the mean as midpoint value might be a result of a general tendency to rate toward the middle. It should be mentioned, too, that the mean playfulness ratings of kindergartners showed the same pattern and thus supports an interpretation of middle-of-the road traits. When the percentage distributions of the ratings are examined, the symmetrical curve shape is further evidence of the trend toward the average.

While the findings may be called acceptable if viewed in the study's own framework and in that of the previous investigations, some questions arise if the findings are compared to others in the field. Wallach and Kogan (1965) who also used ratings in the assessment of classroom nehavior came to the conclusion that a five-point rating scale did not allow for fine enough distinctions. They extended their numerical values to nine while retaining descriptive labels for five points only. The range of means obtained ranged from a low of 2.74 for attention-getting in high-intelligent:low-creative girls to a high of 7.45 in concentration on school work in high-intelligent:high-creative girls. Of course, the study used extreme groups and the need for finer differentiations did not seem to be called for in the present investigations. It would seem worthwhile at some point in our investigations to use one of the other scale of the Wallach and Kogan study as an external criterion for the validity of the PF-nonPF norms. At a later date it is also hoped to use extreme groups on the PF-nonPF dimension and it is, of course, hoped that the group means would be clearly differentiated.

Another angle that needs to be checked out is possible overlapping of behavior. The emerging of two distinctly different factors might give support to this interpretation. Evidence from other studies would be the findings by Holland (1959) of high intercorrelations among teacher-rated variables of originality, popularity, drive to achieve, and physical vigor.

Validity

The question of whether the dimension of PF-nonPF is a psychologically meaningful concept has been answered satisfactorily as far as this phase of the investigation is concerned. Our primary concern at the moment must be internal consistency. Therefore our approach was to establish first logical (empirical) validity based on consensual data as well as the internal check built into the testing situation, namely, the teacher-ranking, which could be considered a concurrent measure.

It is, of course, planned to strengthen the validation of the behavior which signified playfulness by correlating it with measures of divergent thinking, perceptual measures such as Witkin's (1954) Embedded

Figures Test, and finally, other personality measures.

The validity of calling the behavior described playfulness is going to be further checked by a supplementary investigation, which uses a different instrument -- a trait check-list, and a different setting -- community agency centers, for data collection.

Reliability

The most obvious and encouraging conclusions from the reliability data is the consistency with which teachers see their students as playful or nonplayful over time. The most striking and disturbing finding is that two teachers looking at the same students are not able to agree to any satisfactory extent on the degree of PF-nonPF in their students.

Looking more closely first at the test-retest data, it might be possible to conjecture that students change less in their manifestation of PF-nonPF over time with the same teacher than they do from subject to subject at any given same time. This might to some extent also explain the disappointing interrater reliability. Yet, this interpretation is contradicted by the low rank correlations of two pairs of teachers, one in the same subject matter, namely social studies, looking at the class at the same time, and the other, a pairing of English and Modern Languages which in a ranking of ratings in different subject matter areas are in second and third place, respectively. Additional information such as the student teacher's statement that she saw "the students mostly from the back" might provide some explanation for the discrepancy in one case. Of course, with only two pairs of teachers as a sample, it might conceivably have been possible that these were two teachers with widely divergent views of PF-nonPF. Less of a speculation is, however, the fact that this class was extremely homogeneous and therefore not representative of the general sample.

In the HS setting, it is extremely rare to get a population moving in the same track unless it is in a private school as was the case with the sample here. No such difficulty was encountered in the earlier study with kindergartners where two teachers, of equal training, were with the children all the time. The range of interrater reliability on the PF-scales, K-format, was from .66 to .83, perfectly acceptable for personality traits. It may, of course, be claimed that spontaneous behavior in a kindergartner is more easily observed than at the adolescent level, where the classroom calls for more structure and therefore greater inhibition of spontaneous behavior.

In view of these shortcomings of ratings in general and ratings by teachers in particular (Wallach and Kogan 1965; Holland 1959), the major study will assess inter-rater reliability by using two "detached" observers. Several other investigators favor this approach and also argue that it undercuts the halo effect.

Halo

While rater bias in the form of "halo" cannot be altogether eliminated, the first step to minimize this carry-over was in the procedural format of the ratings themselves. The results seem to bear out the usefulness of the separate trait rating sheets when only the correlation coefficients are examined. The different picture presented by the factorial analysis can be understood in the context of the "academic" overtone of one PF-factor. Other investigators, like Holland, report findings similarly confounded with achievement which, naturally, is uppermost in a teacher's perception of the student and might color his evaluation of him in what the teacher considers related areas.

Having two uninvolved observers as an additional check will, of course, also throw further light on whether a true relationship exists or whether it is mainly due to the "halo" effect.

Influence of Secondary Variables

Age and Grade

That age and grade showed no significant correlations with the PF-nonPF dimensions was an unexpected finding. In theoretical formulations of molar behavior such as spontaneity and sense of humor and laughter (Erikson 1963; Berlyne 1966; Piaget 1945; Piddington 1963) developmental differences are stressed as a result of differential functioning in the emotional or cognitive area, or their combination in the affective-cognitive domain. Research studies (Cunningham 1965; Elkind 1966) found older children to be more flexible when populations ranged from age seven to twelve or when elementary school children were compared with adolescents. On the other hand, in an investigation of ethnic attitudes (Wilson 1963) with a population spanning early to late adolescence, it was found that stability of this trait appears to begin early in adolescence.

The clue to an understanding of what, upon first inspection, appears to be a rather startling finding might lie in the age span used. On an empirical basis, it might well appear that the thirteen-year-old is to the teacher the more playful, or conversely, the seventeen-year-old to the investigator. However, if playfulness is a combination of affective-cognitive behavior and if we use Piaget's developmental framework, then the eleven-plus youngster is not necessarily that different from the seventeen-year-old. Such an interpretation calls for a rethinking of theoretical assumptions that cast the early adolescent as a very different individual from the late adolescent in developmental areas other than physical. Even recent studies by Mussen and Jones (1958) do not present such an unequivocal picture, at least on surface traits. Only analysis through projective tests suggested different motivational sources for what appeared to be similar behavior.

While, of course, some weight must also be attached to the purely statistical interpretation of the teacher's frame of reference being limited to her own group and not taking into account an absolute standard, a further developmental check on comparisons of early vs. late adolescents seems very much indicated.

Sex

While there are again no significant differences between boys and girls in the manifestation of PF-nonPF dimensions, the direction of the correlations favors the boys. This finding is in agreement with studies on impulsivity and rigidity (Sutton-Smith 1961; Cunningham 1965) which show girls high on impulsivity to be more masculine in their play preference and boys better able to overcome set.

Looking more closely at the spread of ratings, it is interesting to note the greater variability for boys. It should appear that girls show their conformity here. This general tendency on the part of the girls might further be accentuated by the classroom setting.

These findings are neither new nor startling but they do add support to the validity of the behavior measured.

Subject Matter Area

Although no specific hypotheses had been put forward about the influence of subject matter area on PF-nonPF ratings, the results showing significant differences did not come altogether as a surprise. Of course what must be borne in mind is also the low interrater reliability and therefore inconsistency from one teacher to the other might be as much of a factor as the subject the teacher taught. It might be worthwhile therefore to have independent observers follow the same students to different classes as was actually done by the investigator in her preliminary work, though without precise quantitative evaluation.

Various investigators have attempted to assess interaction in the classroom and its effect on learning (Medley and Mitzel 1963; Amidon and Flanders 1961, 1963). The observation schedule and record (OSCAR) developed by Medley and Mitzel used two observers and established reliabilities on fourteen dimensions which range from .605 for autonomous social grouping to .916 for manifest teacher-hostility. It is not surprising that the higher coefficients are obtained for negative behavior which is at once more conspicuous and less equivocal. It is also interesting to note that three orthogonal factors emerged from the instrument which were labeled Emotional Climate, Verbal Emphasis, and Social Structure. Another widely used instrument is Flanders' interaction analysis of teacher behavior into indirect or direct influence. Some of the studies using the interaction analysis (Amidon and Simon 1965) suggest that the indirect method encourages discovery, that it has differential effects among dependent and independent students, that it enhances teacher effectiveness and leads to greater pupil growth. While there is little reason to challenge the underlying assumptions, the instrument itself invites a great deal of subjective interpretation and some of the results are not clearcut. A tighter approach was used by Macdonald and Zaret (1967) in their classification of verbal behavior in the classroom according to an open and closed process along a continuum. Their findings with a small sample suggested a clearcut differentiation in teaching style paired with learning results. This type of analysis is lengthy and time-consuming but might be worthwhile to apply with a selected sample.

In a more speculative vein one might contend that teacher personality expressed itself in the choice of subject matter area taught and by assessing T-personality, some of the different climates created for PF-nonPF in the high school classroom could be analyzed.

Urban vs. Suburban School

Since the population samples from the urban and suburban schools were comparable, the fact that no differences were found in the ratings as a result of different types of schools points to a possible standardization of both teacher outlook and student behavior. The finding may therefore suggest a conformity that runs across school systems that might either be considered deplorable or encouraging, depending on the point of view and the variables investigated. What it suggests for further work in the study of PF-nonPF is a look at a rural school and such atypical schools as those for the gifted, both academically and in the creative arts.

V - SUMMARY AND CONCLUSIONS

The specific aims of the present project were to establish criteria for playfulness and nonplayfulness in adolescents and to develop measuring instruments for these behavior dimensions. It was hypothesized that teachers were able to rate adolescents on these traits as manifested in the HS classroom. As a corrollary hypothesis, the dimensionality of PF-nonPF was tested.

The results may be summarized as follows:

1. A content analysis of questionnaires completed by 115 JHS and HS teachers, counter-checked with observational criteria obtained by the chief investigator and fellow-psychologists produced a rating instrument whose trait descriptions covered physical, emotional, social and cognitive manifestations of PF-nonPF in adolescents in a HS classroom. Nonplayfulness emerged as a complementary dimension on a continuum.
2. Ratings on the ten subscales of the PF-nonPF instrument showed satisfactory reliability and validity coefficients and thereby attested to the measureability of the behavior.
3. Two distinct factors emerged and, on the basis of their loadings on component traits of PF-nonPF, were labeled "social-emotional" and "academic" PF-nonPF.

The sample for the formulation of the rating scale were 115 JHS and HS teachers drawn from New York City and suburban schools, both public and private. They represented fifteen subject matter areas.

Four psychologists in addition to the investigator were used in counter-checking the teachers' criteria.

The standardization sample for the instrument consisted of 610 JHS and HS students from seven New York City and two suburban schools. The subjects were 338 boys and 272 girls from grades nine through twelve. Age range was from 13.1 to 19.3 years. Four achievement levels--above average, average, below average, and heterogenous grouping--were represented. Twenty-two teachers rated the subjects in twenty-seven different class groups ranging from 6 members to 36 members.

The measuring instrument in the first phase of the investigation, namely, formulation of instrument, was an open-ended questionnaire supplemented by interview data.

The measure of evaluation for the second phase, namely, standardization of the instrument, was a five-point rating scale for playfulness-nonplayfulness, consisting of ten subscales, and two unrelated questions.

The teachers rated each student in their class on the ten subscales making up the PF-nonPF dimensions, namely, physically on the move-physically rigid; physically alert-physically apathetic; enthusiastic-discouraged; relaxed (spontaneous)-tense (constricted); fun-loving-humorless: accepting in wit-hostile in wit; group-oriented-self-oriented; friendly-rejecting; intellectually alive-intellectually stagnant; erratic-conscientous. Ratings on ambitious (achievement-oriented) vs. indifferent and beautiful (handsome) vs. plain (unattractive) were also obtained. After an interval of from four to six weeks, another set of ratings was obtained and at that time the teachers were also asked to rank the students according to their own conceptualization of PF-nonPF.

From the correlational matrix of 35 variables for the total sample, reliability coefficients were obtained for the stability and internal consistency of the separate A and B Scales as well as for the total scale. Uncorrected split-half reliabilities between A and B Scales (quantity and quality dimensions) were .84 on test and .86 on retest. With Spearman-Brown correction, these read .91 on test and .92 on retest. The coefficient of stability over time was .82 for total test. Item reliabilities on test ranged from .36 to .86, with a mean of .68. Item stability from test to retest ranged from .59 to .68, with a mean of .64. The Kuder-Richardson estimate of reliability for test was .87 and for retest .90.

Validity data were drawn from the empirical validity of teacher-agreement on traits ranging from 23% to 2%, and from the correlations of ratings with teachers' ranking, namely, .69 on test and .76 on retest, with mean item validities of .47 on test and .56 on retest.

An examination of the correlations of Scales VI and VII, which tapped achievement-orientation and physical attractiveness, respectively, provided a test for a possible "halo." The mean correlation of the ringer items with total score was .20. Correlations between achievement-orientation and teacher-ranking on PF-nonPF was zero, and between physical attractiveness and teacher-ranking .20.

The dimensionality of the scale was explored by a principal components factor analysis of sixteen variables consisting of age, sex and grade level of student, sex of teacher, and the test ratings on the ten playful-nonplayful traits as well as on achievement orientation and physical attractiveness. These were then rotated by varimax to four factors, using the best fit approach. Two distinct and psychologically meaningful factors emerged for the PF-nonPF dimension. The first factor is made up of physical mobility, spontaneous joy, humor, sociability, and play, and its complementary nonplayful dimensions of physical rigidity, tenseness, lack of humor, rejecting behavior, and conscientiousness. The second factor consists of physical alertness, enthusiasm, intellectual curiosity, and the ringer question assessing achievement-orientation with the complementary behaviors stated as physical apathy, discouragement, intellectual stagnation and indifference.

Inter-rater reliability coefficients were obtained for a subsample of 158 students and ranged from .52 for IIB to .28 for IVB, denoting enthusiasm and friendliness, respectively, with a mean of .38. It was also possible to compute rank correlations for the evaluation of two classes, which were .34 and .27 respectively, neither of them significantly different from zero at the .05 level.

No significant correlations were obtained between the PF-nonPF dimensions and the sex, age and grade of the student, or sex of the teacher. The data from the correlational matrix were supported by the results of the factor analysis which showed age and grade of student to be a separate factor, and sex of teacher and student also a distinct factor, but neither of them related to the PF-nonPF dimensions. Both in the correlations and in the factor analysis, a slight negative direction could be observed with respect to sex, which, indicates that if there is a difference it would favor the boys.

Two one-way analyses of variance assessed the influence of subject matter area and type of school on the PF-nonPF ratings. Except for scales IB and IIA, differences among subject matter areas were significant at the .01 level on all scales. The difference was significant at the .025 level for IB, but not significant for IIA, which is the enthusiastic-discouraged dimension. A two-way analysis of variance which aimed to assess subject matter area in interaction with length of acquaintance (exposure of student to teacher) found the same pattern, though, because of the smaller sample involved, some scales failed to reach significance.

The data from the comparison of urban vs. suburban sample showed no significant differences in ratings in connection with type of school.

The two-way analysis of variance adding length of exposure to subject matter area showed no significant differences for ratings on Scales IB, IIB, IIIA, and IIIB. It might be worth noting that except for the scale on physical alertness-physical apathy (IB), the dimensions which were not influenced by the interaction are all more heavily loaded on the emotional-social factor of PF-nonPF, which is also interpreted as situation-spanning.

The following conclusions may be drawn from the data:

1. Playfulness and nonplayfulness are identifiable in adolescents in the HS classroom and are complementary ends on a rating scale continuum.
2. HS students can be rated by their teachers on PF-nonPF and can be differentiated from one another by numerical rating and ranking. The frequency distributions obtained suggest approximations to the normal distribution curve.

When there is a skew it is in the positive direction.

3. Teachers are able to make a reliable assessment of PF-nonPF behavior in their students as it affects the component traits and thereby attest to the internal consistency of the behavior formulated.
4. Teachers also see their students as relatively stable in PF-nonPF characteristics over a period of one to two months.
5. Since two teachers looking at the same students at the same time only reach minimal agreement on their PF-nonPF characteristics, further work on briefing procedures and sampling is necessary to pinpoint the weakness of the measure as regards interrater reliability.
6. When the rating teachers' conceptualization of PF-nonPF was correlated with the PF-nonPF traits based on the conceptualization of the sample of teachers on which the instrument was formulated, satisfactory agreement was obtained, thereby establishing a validity check across a population of teachers.
7. The emergence of two distinct playfulness-nonplayfulness patterns in the HS classroom suggests an interpretation that might consider the one teacher-approved type of playfulness (academic), especially in view of its high loading on achievement-orientation. It might also be conjectured that the combination of physical alertness, enthusiasm, intellectual curiosity are traits more specific to the school situation. Conversely, physical mobility, spontaneous joy, humor and sociability are not necessarily confined to the classroom,

and, may, in some teachers' eyes, even be frowned upon. Since PF-nonPF in the school is only one aspect of the adolescent's behavior, it might be worthwhile to investigate these situation-specific and situation-spanning traits separately. In particular, it might be worthwhile to follow up the relationship of social-emotional and academic playfulness to achievement level and application to learning in general.

8. The evidence that age and sex are not related to PF-nonPF characteristics was contrary to expectations, both on the part of the investigator and of the teachers. However, while the investigator hypothesized a direct, positive relationship, the teachers conjectured an inverse relationship between age and PF-nonPF. As to sex, both investigator and teachers saw boys as more playful, but this assumption was only borne out by the direction of the correlation but not the size. The breakdown by subsamples and specific traits does suggest, however, one or the other trait may be more differentiating between the sexes. An examination of profiles of extreme cases might give further clues to any differences that might exist.

9. The findings that subject matter area influences PF-nonPF assessment of students has important educational implications. The questions that need to be further explored are whether different subjects allow for more or less play or whether different teachers see more or less of play in their students. In view of the low interrater reliability, this

question also might throw some light on the factors responsible for the low interrater reliability.

10. The evidence of minimal rater bias as shown in the difference in the size of correlation coefficients between PF-nonPF scales and total test score, and the "ringer" questions and total test score is reassuring for the soundness of the instrument in this respect. However, the heavy loading of factor 2 on achievement-orientation calls for a reappraisal of this question as a ringer. The finding itself suggests further research into the connection of academic playfulness and achievement.

Since the present study is part of a long-term investigation to establish criteria of playfulness-nonplayfulness at different developmental levels and then test the relationship of playfulness-nonplayfulness first to cognitive style, especially divergent thinking, as well as to background and environmental variables, the next phase of the research is planned to work with the present instrument in the high school setting.

Although satisfactory reliability and validity have been established and the behavior seems to be meaningful to the situation applied, a supplementary investigation following up the findings of situation-spanning vs. situation-specific traits of playfulness-nonplayfulness in a leisure-type setting has also been proposed. A further check on the dimension of non-playfulness has also been incorporated into the design of this particular study. While in the investigation just completed, nonplayfulness emerged as a complementary trait to playfulness, some comments by raters suggest a relationship of proportionality rather than opposites. A different presentation of the instrument is, therefore, part of the follow-up.

In a larger framework, the present study is seen to contribute to an assessment of motivational variables and/or personality attributes affecting cognitive style. Playfulness was first identified and measured as a quality of play in kindergartners and was hypothesized to become, at later developmental stages, a quality of the player. At the kindergarten level significant relationships were found to divergent thinking tasks. It is considered important, therefore, to continue identifying, measuring and, if and where necessary, encouraging that quality in elementary school children, adolescents and last, but not least, also adults.

REFERENCES

Amidon, E., & Flanders, N. A. The effects of direct and indirect teacher influence on dependent-prone students learning geometry. Journal of Educational Psychology, 1961, 52, 286-291.

Amidon, E., & Flanders, N. A. The role of the teacher in the classroom. Minneapolis: Paul S. Amidon and Associates, 1963.

Amidon, E., & Simon, A. Implications for teacher education of interaction analysis research in student teaching. Paper presented at the American Educational Research Association meetings, Chicago, February, 1965.

Berlyne, D. E. Laughter, humor and play. In G. Lindzey & E. Aronson (eds.), Handbook of social psychology. New York: Addison-Wesley, 1967, in press.

Cunningham, J. D. Einstellung rigidity in children. Journal of Experimental Child Psychology, 1965, 2, 237-247.

Denny, T., Starks, D., & Feldhusen, J. Predictions of divergent thinking and creative performance over a four year period: A longitudinal study. Paper presented at the Annual Convention of the American Psychological Association, Washington, D. C., September, 1967.

Elkind, D. Conceptual orientation shifts in children and adolescents. Child Development, 1966, 37, 493-498.

Emmerich, W. Continuity and stability in early social development. Child Development, 1964, 35, 311-332.

Erikson, E. H. Childhood and society. (2nd ed.) New York: Norton, 1963.

Furst, Norma F. The multiple languages of the classroom: A further analysis and synthesis of meanings communicated in high school teaching. Unpublished doctoral dissertation, Temple University, 1967.

Getzels, J. W., & Jackson, P. W. Creativity and intelligence. New York: Wiley, 1962.

Guilford, J. P. Structure of intellect. Psychological Bulletin, 1956, 53, 267-293.

Guilford, J. P. Psychometric methods. (2nd ed.) New York: McGraw-Hill, 1954.

Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart and Winston, 1963.

Holland, J. L. Some limitations of teacher ratings as predictors of creativity. Journal of Educational Psychology, 1959, 50, 219-223.

Kagan, J. Reflection-Impulsivity: The generality and dynamics of conceptual tempo. Journal of Abnormal Psychology, 1966, 71, 17-24.

Kagan, J. Personal communication. March, 1967.

Kagan, J., & Moss, H. A. Birth to maturity. New York: Wiley, 1962.

Lieberman, J. Nina. Playfulness and divergent thinking: An investigation of their relationship at the kindergarten level. Unpublished doctoral dissertation, Columbia University, 1964.

Lieberman, J. Nina. The relationship between playfulness and divergent thinking at the kindergarten level. Journal of Genetic Psychology, 1965, 107, 219-224.

Lieberman, J. Nina. Playfulness: An attempt to conceptualize a quality of play and of the player. Psychological Reports, 1966, 19, 1278.

MacDonald, J., & Zaret, E. A study of openness in classroom interaction. Paper presented at the American Educational Research Association, New York, February 1967.

McReynolds, P. Toward a theory of fun. Paper presented at the American Psychological Association, Los Angeles, September 1964.

Medley, D., & Mitzel, H. Some behavioral correlates of teacher effectiveness. Journal of Educational Psychology, 1959, 50, 239-246.

Mussen, P. H., & Jones, M. C. Self-conceptions, motivations, and interpersonal attitudes of early-and late-maturing girls. Child Development, 1958, 29, 491-501.

Piaget, J. Play, dreams and imitation in childhood. London: Routledge & Kegan Paul, 1945.

Piddington, R. The psychology of laughter: From Plato to Freud. New York: Gamut Press, 1963.

Rogers, C. R. Toward a theory of creativity. In H. H. Anderson (Ed.) Creativity and its cultivation. New York: Harper & Brothers, 1959. Pp. 69-82.

Sutton-Smith, B. Piaget on play: A critique. Psychological Review, 1966, 173, 104-110.

Sutton-Smith, B. & Rosenberg, B. G. Peer perceptions of impulsive behavior. Merrill Palmer Quarterly, 1961, 7, 233-238.

Thorndike, R. L. Some methodological issues in the study of creativity. In: Proceedings of the 1962 invitational conference on testing problems. Princeton, N. J.: Educational Testing Service, 1963, Pp. 40-54.

Torrance, E. P. Guiding creative talent. Englewood Cliffs, N.J.: Prentice-Hall, 1962.

Wallach, M. A., & Kogan, N. Modes of thinking in young children: A study of the creativity-intelligence distinction. New York: Holt, Rinehart & Winston, 1965.

Wilson, W. C. The development of ethnic attitudes in adolescence. Child Development, 1963, 34, 247-256.

Witkin, H. A. Personality through perception, an experimental and clinical study. New York: Harper & Brothers, 1954.

APPENDIX A

TEACHER QUESTIONNAIRE

Rating Instructions

Dear Teacher:

As you look at a class, you can usually discern individual differences in the behavior of your students.

In our investigation we are interested in a trait called "playfulness," and we would appreciate your cooperation in helping us define this behavior in high-school students.

Attached you will find

1. a Questionnaire in which you are asked to give us your definition of the concept as well as behavioral indices of it in the classroom;
2. a Rating Scale (provisional) for playfulness and a blank sheet for your comments on scales I through V. We are particularly interested to know whether these traits are observable in the classroom, how they agree and differ with your assessment of playfulness and why. Please feel completely free to give us your frank comments on these behavior dimensions. Your constructive criticism will help us to reformulate the scale into a more valid instrument.

Please do not look at the Rating Scale before answering the Questionnaire.

Thank you for your assistance in this project.

Sincerely,

J. Nina Lieberman

J. Nina Lieberman
Assistant Professor
of Education

Name:

School:

Subject Area:

HT - 1

J. N. Lieberman

Brooklyn College

QUESTIONNAIRE

1. Write down any ideas or concepts that are suggested by the word "playfulness."

2. State any behavior that would describe a playful teenager. Be as detailed and specific as possible.

3. How would you identify a "playful student" in your classroom? Give specific behavioral traits or incidents.

Name:
School:
Subject Area:

HT - 1
J. N. Lieberman
Brooklyn College

4. Write down any ideas or concepts that are suggested by the word "nonplayfulness."

5. State any behavior that would describe a nonplayful teenager. Be as detailed and specific as possible.

6. How would you identify a "nonplayful student" in your classroom? Give specific behavioral traits or incidents.

Name:
School:
Subject Area:

HT - 1
J. N. Lieberman
Brooklyn College

7. In your opinion, does playfulness influence classroom learning?

Yes No

If yes, How?

8. In your opinion, does nonplayfulness influence classroom learning?

Yes No

If yes, How?

Form A-1

RATING SCALES

1. A. How often does the boy (girl) manifest spontaneous physical movement and activity in class?

This behavior would include facial animation like a happy or questioning look, gestures, dance-like steps, and other rhythmic movements of the whole body or parts of the body which could be judged as a fairly clear indication of exuberance.

Very often	Often	Occasionally	Rarely	Very Rarely
5	4	3	2	1

B. How is his (her) general motor coordination?

Excellent	Very Good	Good	Fair	Poor
5	4	3	2	1

II. A. How often does the boy (girl) show joy in or during his classroom activities?

This may be judged by facial expression such as smiling, by verbal expressions such as saying, "this is great," or other currently used (popular) phrases of approval. Other behavioral indicators would be repetition of activity, or resumption of activity, with clear evidence of enjoyment.

Very often	Often	Occasionally	Rarely	Very Rarely
5	4	3	2	1

B. With what freedom of expression does he (she) show his joy?

This may be judged by the cadence and modulation of a chuckle and/or verbalizations.

Very High	High	Moderate	Some	Little
5	4	3	2	1

III. A. How often does the boy (girl) show a sense of humor?

By sense of humor is meant punning, seeing far-out analogies and an ability to make a joke and be the butt of one.

Very Often	Often	Occasionally	Rarely	Very Rarely
5	4	3	2	1

B. What degree of hostility is mixed into the humor?

This may be judged by the subtlety of a tease, or an admixture of gentleness and mischief.

Very High	High	Moderate	Some	Little
5	4	3	2	1

IV. A. To what extent does the boy (girl) show flexibility in interaction with the surrounding group structure?

This may be judged by role changes, i.e., from leader to follower, from information-seeker to information-facilitator in a given group setting, and by the youngster moving in and out of these groups after having been an integral part of them.

Very Often	Often	Occasionally	Rarely	Very Rarely
5	4	3	2	1

B. With what degree of ease does he (she) handle the changes in interaction?

This may be judged by ready acceptance of the new roles, by lack of distress over a change in the group situation, including also an ability to amuse himself (herself) if left solitary after peer interaction.

Very high	high	Moderate	Some	Little
5	4	3	2	1

V. A. How often does the boy (girl) show spontaneity in intellectual tasks?

This may be judged by flow of ideas, ability to change set and toy with concepts and relationships, finding remote connections or unusual solutions in tackling set problems, as well as performance in free discussions and dramatic presentations.

Very Often	Often	Occasionally	Rarely	Very Rarely
5	4	3	2	1

B. What degree of imagination does he (she) show in these activities?

This may be judged by the original quality of approaching a theme or experiment, by figures of speech used in discussions, and by depth and breadth of dramatic presentations.

Very High	High	Moderate	Some	Little
5	4	3	2	1

VI. How bright is the boy (girl)?

This is your estimate of the youngster's intelligence based on achievement or inferred potential.

Extremely Bright	Bright	Average	Moderately Bright	Not too Bright
5	4	3	2	1

VII. How attractive is the boy (girl)?

This is your evaluation of the youngster's physical appeal.

Beautiful	Very Attractive	Nice - Looking	Passable in looks and appearance	Somewhat homely and Unattractive
5	4	3	2	1

Name:

School:

Subject Area

Playfulness Study (A-1)
J. Nina Lieberman
Dept. of Education
Brooklyn College

COMMENTS ON SCALES IA THROUGH VB OF PLAYFULNESS SCALE (ADOLESCENT FORM)

APPENDIX B

PLAYFULNESS-NONPLAYFULNESS SCALE
Form (A)

RATING INSTRUCTIONS

As you look at adolescents in a classroom setting, you realize that they differ in the way they move about, address themselves to their tasks, and interact with peers and teachers.

In this study we are trying to assess how much spontaneity can be found in the behavior of high-school students in the classroom. Also, how cheerful and how "full-of-the-devil" these youngsters are.

Attached you will find a rating measure made up of five scales which refer to a student's behavior in class. You will note that each of the five scales or questions has two parts. Part A of the question aims at measuring the frequency or quantity of the trait; Part B tries to assess the quality of the trait shown. For example, "how consistently does the student show a sense of fun?" would be the quantity of the trait, "and how much is wit and subtlety a part of his sense of humor?" would be the quality of the trait.

We hope that we shall have your cooperation in this work and that you will find it possible and worthwhile to look at the students in your classroom along the traits suggested in the rating scales and give us your evaluation of them.

We are also interested in finding out what your impression is of the student's achievement orientation and physical attractiveness and would like you to give us your estimate of these as well.

When you rate the students, you will, of course, want to compare them with one another as well as keep in mind a general standard for these traits in adolescents in the high-school setting.

It is easier and better to rate all students first on one trait or question and then do the same for the six others. The rating scales have, therefore, been set up for one trait per page.

PLEASE PUT DOWN THE FIGURE THAT BEST INDICATES YOUR EVALUATION OF THE STUDENT'S PRESENT STANDING.

A PROFILE IS GIVEN AT THE EXTREME ENDS OF EACH SCALE AS AN AID IN MAKING YOUR RATING. THE SCALE IS TO BE REGARDED AS A CONTINUUM AND THE IN-BETWEEN NUMBERS SHOULD BE USED TO INDICATE DEGREES IN FREQUENCY AND INTENSITY.

Any comments about the content or form of the questions, or about any difficulties that you may have in answering them, will be welcomed.

Thank you for your help in this study.

[illegible]-NONPT (A)

Name of Student:
Grade (or Level):
Subject:
Teacher:
School:

Scale
IA. How consistently does the student show spontaneous physical movement and activity in class?

Physically on the move	Physically rigid
5 4 3	2 1
This is the student who moves around a lot, likes to change his seat has trouble settling down, fidgets with things, mischievously throws objects.	This is the student who sits stiffly, with a tense facial expression, and a rigid manner.

Scale
IB What degree of energy does the student show in physical activity?

Physically alert	Physically apathetic
5 4 3	2 1
This is the student who has an animated and alert facial expression, waves his hand to be recognized, uses gestures freely to underline a point, nods in response to teacher's points.	This is the student who slumps in his seat, looks sluggish. and sleepy, yawns, stares into space.

PHENOTYPE (A)

Name of Student:
Grade (or Level):
Subject:
Teacher:
School:

SCALE
IIA

How consistently does the student show enthusiasm during classroom activities?

Enthusiastic — Discouraged

5 — 4 — 3 — 2 — 1

This is the student who is eager and enthusiastic in his approach to work, optimistic and high-spirited.	This is the student who needs reassurance, is unhappy and sullen, gets easily discouraged, is unsure of himself.

SCALE
IIB

With what ease (freedom) does the student show joy?

Relaxed (Spontaneous) — Tense (Constricted)

5 — 4 — 3 — 2 — 1

This is the student who is relaxed, boisterous, chuckles and laughs, can express feelings freely, sometimes unpredictably.	This is the student who is tense, quiet, rarely laughs lacks spontaneity, stereotype in reactions (you just know what he is or is not going to do next)

PROFILE (A)

Name of Student:
Grade (or Level):
Subject:
Teacher:
School:

SCALE
IIIA

How consistently does the student show a sense of fun (humor) in class?

Fun-loving — Humorless

5 — 4 — 3 — 2 — 1

This is the student
who is the entertainer,
who constantly makes jokes,
enjoys horseplay,
clowns,
engages in cross-sex
teasing.

This is the student
who becomes irritable
in a fun situation,
who is anxious to get
back to the "real
business"--the lesson,
who fails to see the
funny side of
situations.

SCALE
IIIB

How much is wit and subtlety a part of his sense of humor in class?

Accepting in wit — Hostile in wit

5 — 4 — 3 — 2 — 1

This is the student
who recognizes, and searches for,
the humor in situations,
can take teasing and teases others,
including the teacher,
who uses wit in puns, off-beat
comparisons, and sometimes
slightly off-color remarks.

This is the student
who laughs at the dis-
comfort of others,
gets angry when he
himself is the butt
of a joke,
hits back with insults
when teased.

PF-NONPF (A)

Name of Student:
Grade (or Level):
Subject:
Teacher:
School:

SCALE
IVA

How consistently is the student engaged in interaction with peers in class?

Group-oriented — Self-oriented

5 — 4 — 3 — 2 — 1

Group-oriented	Self-oriented
This is the student who is busy passing notes, talking to neighbors, seeking attention also by pushing and shoving, and calling out in class.	This is the student who keeps to himself, "a loner", does not respond to classmates, and does not, on his own, seek association with them.

SCALE
IVB

What is the tone or quality of the involvement with peers in class?

Friendly — Rejecting

5 — 4 — 3 — 2 — 1

Friendly	Rejecting
This is the student who is outgoing, friendly, able to move from one group to another.	This is the student who gets easily hurt, is on the defensive with others, wants to hurt others uncooperative.

PF-NONPF (A)

Name of Student:
Grade (or Level):
Subject:
Teacher:
School:

SCALE
VA

How consistently does the student show spontaneity in intellectual tasks in class?

Intellectually alive				Intellectually stagnant
5	4	3	2	1

This is the student
who is curious,
inventive,
volunteers frequently,
introduces relevant
and sometimes far-out
material in questions and
answers.

This is the student
who approaches work in
a routine and mechanical
way,
does not volunteer in
class, and, when called
upon, does not respond,
sometimes daydreaming
or appearing bored.

SCALE
VB

What is the quality of the student's work involvement in class?

Erratic				Conscientious
5	4	3	2	1

This is the student
more concerned with
play than work,
bored with the regular
classroom atmosphere,
and off in his own world,
sometimes asking questions
to disrupt the lesson.

This is the student
who is conscientious,
completes his assignments,
takes his work seriously-
sometimes too seriously.

PF-NOMF (A)

Name of Student:
Grade (or Level):
Subject:
Teacher:
School:

SCALE
I
How achievement-oriented is the student?

Ambitious	Indifferent

5 4 3 2 1

This is the student whose whole attention is on excellence as shown by grades and tests.	This is the student who cares little or nothing about his academic standing in class.

SCALE
VII
How attractive is the student physically?

Beautiful (Handsome)	Plain (Unattractive)

5 4 3 2 1

This is the student who is exceptionally good-looking.	This is the student who is homely and unattractive.

SAMPLE RATING SHEET

PF-NONPF (A)

Grade (or Level):
Subject:
Teacher:
School:
Date:

IA. How consistently does the student show spontaneous physical movement and activity in class?

Physically on the move — Physically rigid

5 4 3 2 1

This is the student who moves around a lot, likes to change his seat, has trouble settling down, fidgets with things, mischievously throws objects.

This is the student who sits stiffly, with a tense facial expression, and a rigid manner.

PLEASE PUT DOWN THE FIGURE THAT BEST
INDICATES THE STUDENTS'S PRESENT STANDING
ACCORDING TO THE SCALE GIVEN ABOVE.

NAME	RATING	NAME	RATING